Konrad-Zuse-Zentrum
für Informationstechnik Berlin

RALF BANISCH[1], NATAŠA DJURDJEVAC CONRAD[1] AND
CHRISTOF SCHÜTTE[1,2]

[1]*Department of Mathematics and Computer Science, Freie Universität Berlin, Germany*

[2]*Zuse Institute Berlin, Germany*

Reactive flows and unproductive cycles for random walks on complex networks

Reactive flows and unproductive cycles for random walks on complex networks

Ralf Banisch, Nataša Djurdjevac Conrad and Christof Schütte

Institut für Mathematik, Freie Universität Berlin, Germany
and
Zuse Institute Berlin, Germany

Abstract

We present a comprehensive theory for analysis and understanding of transition events between an initial set A and a target set B for general ergodic finite-state space Markov chains or jump processes, including random walks on networks as they occur, e.g., in Markov State Modelling in molecular dynamics. The theory allows us to decompose the probability flow generated by transition events between the sets A and B into the productive part that directly flows from A to B through reaction pathways and the unproductive part that runs in loops and is supported on cycles of the underlying network. It applies to random walks on directed networks and nonreversible Markov processes and can be seen as an extension of Transition Path Theory. Information on reaction pathways and unproductive cycles results from the stochastic cycle decomposition of the underlying network which also allows to compute their corresponding weight, thus characterizing completely which structure is used how often in transition events. The new theory is illustrated by an application to a Markov State Model resulting from weakly damped Langevin dynamics where the unproductive cycles are associated with periodic orbits of the underlying Hamiltonian dynamics.

Keywords: Complex networks, molecular transition networks, transition path theory, cycle decomposition, reactive trajectories, Markov State Models

AMS classification: 05C81, 90B15

1 Introduction

Rare but important transition events between long lived states are a key feature of many systems arising in physics, chemistry, biology, etc. Molecular dynamics (MD) simulations allow for analysis and understanding of the dynamical behaviour of molecular systems. However, realistic simulations for large molecular systems in solution on timescales beyond milliseconds are still infeasible even on the most powerful general purpose computers. Rare events require prohibitively long simulations because

the average waiting time between the events is orders of magnitude longer than the timescale of the transition characterising the event itself. Therefore, the straightforward approach via direct numerical simulation of the system until a reasonable number of events has been observed is infeasible for most interesting systems.

We consider rare events in which molecular system under consideration has the ability to go from a initial state given by a set A in its state space (e.g. an initial conformation) to a target state described by another set B (e.g. the target conformation). A is a metastable set and thus transitions from A to B are rare. We want to *characterize* the transitions leading from A into B, that is, we are interested in the statistical properties of the ensemble of *reactive trajectories* that go *directly* from A to B (i.e. start in A without returning to A before going to B). We would like to know where in state space the probability flow generated by reactive trajectories is largest, i.e., which transition channels in state space reactive trajectories prefer and how a typical transition event happens.

In this article, we will discuss the estimation of rare event statistics via *discretization* of the state space of the system under consideration. That is, instead of dealing with the computation of rare events for the original, continuous process using MD trajectory simulation like, e.g., in Transition Path Sampling, we will utilize so-called Markov State Model (MSM) with finite, *discrete* state space. The reason is that for such a discrete model one can numerically compute the probability flow generated by reactive trajectories and the associated transition channels completely *without* simulation. Instead, discrete Transition Path Theory (TPT) [16] shows that based on an MSM one just has to solve finitely many linear equations in order to get the reactive probability flow *everywhere* in state space.

In the standard setting a Markov State Model (MSM) is a Markov chain whose transition matrix is given by the transition probabilities of the original MD process between some subsets of the molecular state space (i.e., molecular conformations) that form the (macro-)states of the MSM. It has been demonstrated that, in many cases of practical relevance ranging from peptide conformation dynamics via protein folding and function to membrane fission, or RNA kinetics, MSM building allows the study of dynamic behavior on long timescales and the analysis and understanding of rare transition events using TPT, see [4, 19, 26, 21, 20, 11] for just some links to the relevant literature.

An MSM can be interpreted as a transition network of the molecular system [18, 17] by taking the MSM (macro-)states, i.e., conformation sets of the underlying molecular system, as the nodes of the network, putting edges where these nodes are connected by positive transition probabilities and consider the random walk on the thus constructed network that is induced by jumping from node to node according to the respective transition probabilities, compare Fig. 1.

In recent years, such network descriptions and the associated random walks have attracted a lot of attention as rather general tools to model, represent, and interpret complex processes arising in many fields of science, ranging from transition networks in molecular research and reaction networks in cellular biology, via logistics, computer and supply networks to phenomena like social networks. The framework of Transition Path Theory (TPT) has been introduced to analyze the statistical properties of the reactive trajectories by which transitions occur between two specific sets of nodes of the network that represent our initial state A and target state B. TPT allows for extraction of the essential features of the dynamics on the network and relates them to the global structure of the network [12, 23].

TPT has been developed in [6, 7, 15, 8] in the context of diffusions, and has

2

been generalized to Markov processes with discrete state space and random walks on networks in [16, 14]. In principle, TPT allows for the analysis of all features of interest regarding the set of all reactive trajectories and its statistical properties: transition rate, probability flow induced by reactive trajectories, reactive flow carried by specific transition pathways, etc. It also permits to compute the list of all important reactive pathways including their weight relative to the overall transition rate, see [16, 14, 25]. However, this approach is limited to direct pathways from A to B that do not include loops. This means, TPT does *not* allow for understanding of the reactive probability flow that is carried by loops in the network. Understanding this "unproductive" flow is important for network design questions (robustness) or for Markov processes out of equilibrium such that an extension of TPT addressing this problem is desirable. For undirected networks (and the associated reversible Markov processes) this extension has been discussed in [3] where it has been demonstrated how to effectively ignore the flow carried in loops. The present work can be viewed as a more general extension that presents a complete theory for solving the problem of characterizing the unproductive flow for undirected and directed networks and associated Markov processes.

In a nutshell, the basic idea is to utilize the stochastic cycle decomposition [9, 5] of the probability flow generated by the random walker, and associate it with TPT aiming at a decomposition of the productive/reactive and unproductive flow between two sets of nodes of the network.

The remainder of this paper is organized as follows. In Section 2 we discuss the relation between Markov processes and networks, and outline the stochastic cycle decomposition. Next, in Section 3 we summarize the main aspects of TPT. Section 4 then shows how to utilize the stochastic cycle decomposition for decomposing the flow into a productive and unproductive part by using the toolset provided by TPT. In Section 5 we apply the tools introduced earlier to analyze random walks in two simple mazes and Langevin dynamics. Finally, some concluding remarks are given in Section 6.

2   Cycle Decomposition of Markov Chains

In all of the following we consider an irreducible aperiodic homogeneous Markov chain $(X_t)_{t\in\mathbb{N}}$ on state space $\mathbb{X} = \{1, \ldots, n\}$ with transition matrix P and transition probability p_{ij} for $i, j \in \mathbb{X}$. We assume that μ denotes the invariant measure of the chain so that $\mu^T = \mu^T P$. The chain is ergodic, i.e., infinitely long trajectories will visit an arbitrary state i with probability μ_i.

P can be the transition matrix of a Markov State Model with n (macro-)states. P can as well be associated with random walks on social networks or the Internet [12], or a reaction network in systems biology, cf. [16].

2.1   Markov chains and random walks on networks

We assume that the chain in general is irreversible, i.e. the detailed balance condition $\mu_i p_{ij} = \mu_j p_{ji}$ is not assumed to be satisfied. Simple examples for such a chain are given by random walks on directed networks: Obviously every chain generates a directed network with node set $\mathbb{X}$ and edge set $E = \{(i, j) : p_{ij} > 0\}$ and edge weights p_{ij}. In turn, a strongly connected graph/network $G = (\mathbb{X}, E)$, where $\mathbb{X}$ is the set of n nodes and E the set of edges of the graph, allows to define an ergodic chain. We denote the adjacency matrix of the network by $(a_{ij})_{i,j\in\mathbb{X}}$ and the out-degree of

a node i by $d_i = \sum_{j \in \mathbb{X}} a_{ij}$. If the network has weighted edges then a_{ij} denotes the weight of the edge $(i, j) \in E$ and $a_{ij} = 0$ if $(i, j) \notin E$. Whenever the network is undirected the adjacency matrix is symmetric, and the associated random walk satisfies detailed balance (i.e., it is reversible). Usually, in network clustering one considers the standard random walk defined on the network, i.e., the Markov chain with one-step transition matrix directly given by the adjacency structure / weights,

$$p_{ij} = \frac{a_{ij}}{d_i}. \tag{1}$$

The associated Markov chain is irreducible and aperiodic, and if the network is undirected and the random walk reversible, it has invariant measure $\mu_i = d_i / \sum_{i \in \mathbb{X}} d_i$, i.e., nodes with high degree are visited often. Random walks where p_{ij} is defined in terms of a_{ij} in a different way [23] have also been considered and connect directed networks to irreversible random walks as well. The following table summarises commonly used network terminology.

Term	Definition
Adjacency matrix $(a_{ij})_{i,j \in \mathbb{X}}$	indicates which nodes are adjacent to which other nodes $$a_{ij} = \begin{cases} w_{i,j} & (i,j) \in E \\ 0 & (i,j) \notin E \end{cases}$$ where $w_{i,j}$ is the weight of the edge (i, j) for weighted graphs or $w_{i,j} = 1$ for unweighted graphs.
Out-degree of a node d_i	is the number of outgoing edges from a node i $d_i = \sum_{j \in \mathbb{X}} a_{ij}$.
In-degree of a node d_i^-	is the number of incoming edges to a node i $d_i^- = \sum_{j \in \mathbb{X}} a_{ji}$.
A source node	is a node with no incoming edges, i.e. $d_i^- = 0$.
A sink node	is a node with no outgoing edges, i.e. $d_i = 0$.
A strongly connected graph	is a graph for which there is a directed path from every node to every other node in the graph.

Molecular transition networks. Fig. 1 shows the transition network that results from MSM building based on MD simulations of 12-alanine, see [18] for details. This network is undirected because the MSM for 12-alanine comes out to be reversible. Obviously, there are several reaction pathways along which the transition process starting in the β-sheet conformation A and ending in the α-helical conformation B

4

will happen. Based on the MSM for 12-alanine, TPT allows to rank them according to their contribution to the overall transtion rate.

However, in general MSM building has to deal with non-reversible MSMs, even if this often is ignored. For example, if the underlying MD simulations are based on Langevin dynamics the MSM will be non-reversible, see [25] for the theoretical background and [20] for the typical handling of it. Furthermore, whenever MD simulations out of equilibrium are concerned the resulting MSM is non-reversible in general even if the underlying MD process may be reversible [28]. In these cases the resulting transiition network is directed.
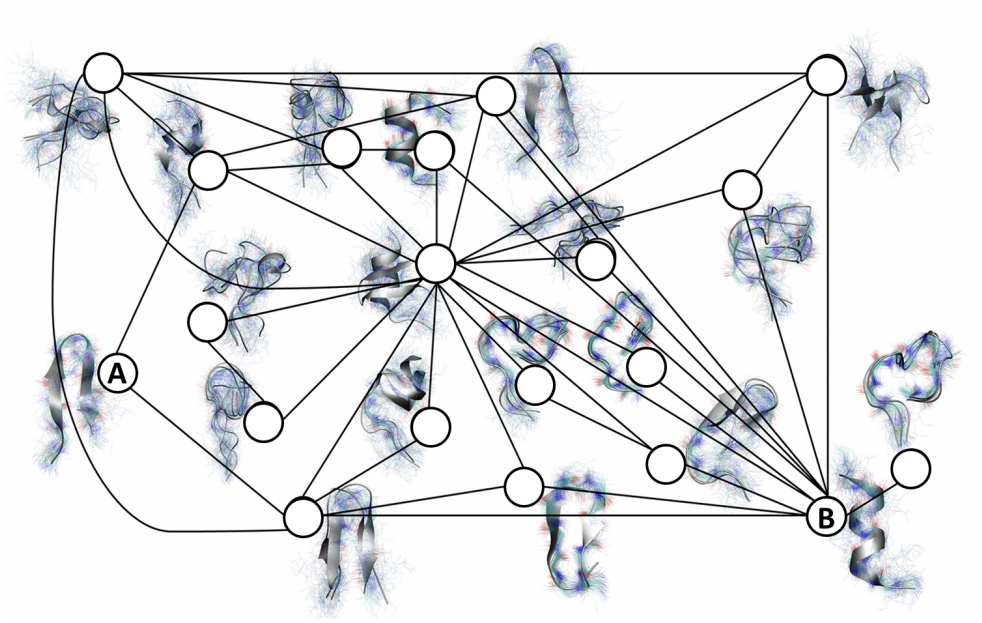

Figure 1: Transition network of 12-alanine as of [18]. The circles indicate the nodes of the network; the associated conformations are shown close to each node. The node marked A is associated with a β-sheet like conformation while B marks an α-helical conformation. If an edge exists then the MSM transition matrix for 12-alanine has a positive transition probability between the respective conformations. The transition probability themselves are omitted.

2.2   Flow decomposition

We give a short account to flow decompositions in terms of cycles. A cycle in the directed graph $G = (\mathbb{X}, E)$ is a closed path $\gamma = (i_1, i_2, \ldots, i_s)$ visiting the nodes $i_1, \ldots, i_s$ in that order, where all edges (i_k, i_{k+1}) and (i_s, i_1) have to be in E. Cycles are called simple if they are non-intersecting, i.e. all i_k are pairwise different. We denote the set of all simple cycles in G by Γ. We write $(i, j) \subset \gamma$ if the edge (i, j) is an edge of γ and their orientations agree. To each cycle $\gamma \in \Gamma$, one can associate a matrix C_γ that encodes which edges are part of γ by

$$C_{\gamma,ij} = \begin{cases} 1, & (i,j) \subset \gamma \\ 0, & \text{otherwise.} \end{cases}$$

5

The chain $(X_t)_{t\in\mathbb{N}}$ induces a flow F_{ij} on the network,

$$F_{ij} = \mu_i p_{ij} = \mathbf{P}[X_{t+1} = j | X_t = i],$$

which simply is the ergodic probability flow through edge (i,j). Because μ is invariant, F is conserved at every node:

$$\sum_i F_{ij} - \sum_i F_{ji} = 0 \quad \forall j \in \mathbb{X}. \tag{2}$$

Equation (2) is typically called *Kirchhoff's loop law* and allows for a decomposition of F into elementary currents along cycles. Different approaches for doing this have been proposed in the literature [24, 29, 1], in general decompositions of this kind depend on a choice of basis or spanning tree of G and are therefore not unique. We focus on the unique *stochastic cycle decomposition* [10, 9, 5] which is of the form

$$F = \sum_{\gamma \in \Gamma} w(\gamma) C_\gamma, \tag{3}$$

where the flow weight $w(\gamma)$ has a stochastic interpretation in terms of the number of times a realization of $(X_t)_t$ actually uses γ; we explain this below. If $\gamma = (i_1, \ldots, i_s)$, then $w(\gamma)$ is given by the explicit formula [9]

$$w(\gamma) = p_{i_1 i_2} \ldots p_{i_{s-1} i_s} p_{i_s i_1} \frac{D(\{i_1, \ldots, i_s\})}{\sum_{i \in \mathbb{X}} D(\{i\})} \tag{4}$$

where $D(\{i_1, \ldots, i_s\})$ is the determinant of the matrix $I - P$ with rows and columns indexed by $\{i_1, \ldots, i_s\}$ deleted. For large state spaces, equation (4) is numerically inappropriate to compute $w(\gamma)$. Another characterization of the cycle weights uses realizations of the underlying Markov chain instead. To this end, consider a realization $(X_t)_{t=1,\ldots,T}$ which defines a path in G with a finite number of self-intersections. Each self-intersection corresponds to a cycle $\gamma \in \Gamma$ that has been passed through by $(X_t)_{t=1,\ldots,T}$. For every $\gamma \in \Gamma$ we define N_T^γ as the number of times $(X_t)_{t=1,\ldots,T}$ passes through γ. Then

$$w(\gamma) = \lim_{T \to \infty} \frac{N_T^\gamma}{T}. \tag{5}$$

Equation 5 will be used later to compute $w(\gamma)$ in numerical examples.

3 Transition Path Theory Revisited

Let A and B be two non-empty, disjoint subsets of the state space $\mathbb{X}$ such that $(A \cup B)^c$ is non-empty. By ergodicity, (X_t) visits both A and B infinitely often and thus oscillates infinitely many times between A and B. We are interested in understanding how these $A \to B$ transitions happen (pathways, rate, etc). In Transition Path Theory (TPT) one views A as a reactant state and B as a product state, such that each $A \to B$ transition is a reaction event, and one asks about the mechanism, rate, etc. of these reaction events. TPT proceeds by pruning out of a long ergodic trajectory of (X_t) the pieces during which it makes a direct transition from A to B (start in A and go to B without detours back to A). These pieces are called the *reactive trajectories* and TPT allows to characterize various statistical properties of these pieces, especially the *reactive probability flow* F^{AB} they generate on state space. F^{AB} is the part of F

which is made up by the reactive trajectories, such that F_{ij}^{AB} is the probability that a reactive trajectory makes a transition from state i to state j. The TPT framework has been developed in [6, 7, 15, 8, 27] in the context of diffusions, and has been generalized to discrete state spaces in [16, 14]. According to [16, 14] the reactive flow between A and B is given by

$$F_{ij}^{AB} = \begin{cases} \mu_i \, q_i^- \, L_{ij} \, q_j^+, & \text{if } i \neq j \\ 0, & \text{otherwise} \end{cases} \qquad (6)$$

where L_{ij} denotes the entries of the rate matrix or *generator* defined by $L = P - \text{Id}$, and q^+ and q^- are the so-called committor functions that are defined as follows: The *forward committor* q_i^+ at state i is defined as the probability that the process starting in $i \in \mathbb{X}$ will reach first B rather than A. Analogously, the *backward committor* q_i^- is defined as the probability that the process arriving in state i came last from A rather than B. The forward and backward committor both satisfy a discrete Dirichlet problem [16, 14]:

$$\begin{cases} \sum_{j \in \mathbb{X}} L_{ij} q_j^+ = 0, & \forall i \in (A \cup B)^c \\ q_i^+ = 0, & \forall i \in A \\ q_i^+ = 1, & \forall i \in B \end{cases} \qquad (7)$$

and

$$\begin{cases} \sum_{j \in \mathbb{X}} \tilde{L}_{ij} q_j^- = 0, & \forall i \in (A \cup B)^c \\ q_i^- = 1, & \forall i \in A \\ q_i^- = 0, & \forall i \in B \end{cases} \qquad (8)$$

Here $\tilde{L} = (\tilde{L}_{ij})$ denotes the generator backward in time given by $\tilde{L}_{ij} = \frac{\mu_j}{\mu_i} L_{ji}$.

Based on the flow one can easily compute the total transition rate of reactive trajectories from A to B by summing F_{ij}^{AB} up either over all edges (ij) that leave A or all edges that go into B, that is

$$k_{AB} = \sum_{i \in A, \, j \in \mathbb{X}} F_{ij}^{AB} = \sum_{i \in \mathbb{X}, \, j \in B} F_{ij}^{AB}.$$

As shown in [16], the transition rate is identical to the expected number of reactive trajectories (i.e. reactive pieces of a long ergodic trajectory) per unit of time. It turns out that k_{AB} can also be expressed by the *effective reactive flow* F^+ which is obtained by removing the reactive trajectories that go from j to i from the ones that go from i to j for every edge (ij), i.e.

$$F_{ij}^+ = \max(0, F_{ij}^{AB} - F_{ji}^{AB}) \qquad (9)$$

such that $k_{AB} = \sum_{i \in A, \, j \in \mathbb{X}} F_{ij}^+$. Once can also compute the probability μ_i^R to find a reactive trajectory at state i, given by

$$\mu_i^R = \begin{cases} q_i^- \, \mu_i q_i^+, & i \in (A \cup B)^c, \\ 0 & \text{otherwise} \end{cases} . \qquad (10)$$

7

4 Reactive and Unproductive Flow

The aim of this section is to develop a decomposition of the probability flow F into the reactive part from a reactant state (subset A of nodes in the network) to a product state (subset B) and the non-reactive or unproductive part. The guiding idea is the concept of Hodge-Helmholtz decompositions of vector fields: Let F be a vector field on some bounded domain $D \subset \mathbb{R}^3$. Then F can always be decomposed into a gradient and a rotation:

$$F = \nabla\Phi + \nabla \times R. \tag{11}$$

The gradient part $\nabla\Phi$ is rotation-free while the rotational part $\nabla \times R$ is divergence-free. Φ and R can be computed by taking the divergence of (11), which yields $\nabla \cdot F = \Delta\Phi$, and the rotation of (11), which yields $\nabla \times F = \nabla \times (\nabla \times R)$. The discrete analogue of vector fields are flows it the network G, which is why we are now looking for a decomposition of the type (11) for flows. We have no differential operators at our disposal, but the rotational part $\nabla \times R$ corresponds to the cyclic flows discussed in section 2.2.

4.1 Simple sink and source

We start with the following toy system to make things conceptually clear: Consider a network with a single point source $A \in \mathbb{X}$, a single point sink $B \in \mathbb{X}$ and a set of strongly connected transient states T. That is $\mathbb{X} = T \cup \{A\} \cup \{B\}$, and

$$p_{ia} = p_{bi} = 0 \quad \forall i \in T, \ \forall a \in A, \ \forall b \in B. \tag{12}$$

Our aim is to characterize the reactive flow which can be reframed as a characterization of information transport from A to B: Start a random walker at A; it will wander around in T for a while and at some point reach B, which we count as one bit of information delivered. We then start the next random walker in A, and so on. This can be modelled by adding a single directed edge $e_{BA} = (B, A)$, see Fig. 2 below. This addition makes the process ergodic and the flow $F(e_{BA})$ will count the number of transitions from A to B, hence the number of bits delivered. Because of assumption (12), a trajectory started in any point $i \in T$ must reach B next (and not A) with probability 1, and it must have come from A (and not B) with probability 1. In other words, $q_i^+ = q_i^- = 1$ for all $x \in T$. As a consequence, $F_{ij} = f_{ij}^{AB}$ for $(i, j) \neq e_{BA}$ and $F(e_{BA}) = k_{AB}$: For this system, the pruning performed by TPT only removes transitions along the edge e_{BA}, all other parts of a generic trajectory are considered reactive. Nevertheless, on its way from A to B, some parts of the trajectory of $(X_t)_t$ will be productive, meaning that they help advancing towards B, while others will be non-productive. Identifying the productive parts of a trajectory is the key definition we have to make. We will do this pathwise:

Consider a single piece of reactive trajectory $p = (A, i_1, \ldots, i_s, B)$ with $i_1, \ldots, i_s \in T$ representing one transition from A to B. After B, the next state has to be A since e_{BA} is the only edge leaving B, so p closes to a cycle $c = (A, i_1, \ldots, i_s, B, A)$. Because the transition from A to B may not be direct, c may have self-intersections and can therefore be decomposed into several simple cycles, exactly one of which contains the edge e_{BA}. We will denote this cycle γ_P. γ_P represents the productive part of c, namely a direct transition from A to B with no detours. All the other simple cycles in the decomposition of c represent the unproductive detours on the way from A to B. This leads to a splitting $\Gamma = \Gamma_P \cup \Gamma_U$ where $\Gamma_P = \{\gamma \in \Gamma : e_{BA} \text{ is an edge of } \gamma\}$

is the set of **productive cycles** which represent transitions from A to B, and $\Gamma_U = \{\gamma \in \Gamma : e_{BA}$ is not an edge of $\gamma\}$ is the set of **unproductive cycles** which represent detours. Observe that cycles in Γ_U must lie completely in the transition region T. In terms of this splitting, (3) becomes

$$F = \sum_{\gamma \in \Gamma_P} w(\gamma) C_\gamma + \sum_{\gamma \in \Gamma_U} w(\gamma) C_\gamma = F_P + F_U. \tag{13}$$

The *reactive flow* or *information flow* F_P is the flow generated by the productive parts of the trajectories, and the *unproductive flow* F_U is the flow generated by the unproductive part. The decomposition $F = F_P + F_U$ is unique and based on the definition of productive and unproductive parts given. Moreover, note that

$$k_{AB} = F(e_{BA}) = F_P(e_{BA}) = \sum_{\gamma \in \Gamma_P} w(\gamma).$$

This is evident, we just need to evaluate (13) on the edge e_{BA}. By the construction above, every reactive trajectory is in one-to-one correspondence to one productive cycle $\gamma \in \Gamma_P$, and the weight $w(\gamma) = \lim_{T \to \infty} \frac{1}{T} N_T(\gamma)$ of this cycle according to (5) indicates how many trajectories per unit time used it as a reaction pathway. Thus, F_P encodes the transport mechanism from A to B.

Example. Consider the network given in Fig. 2, where all edges have weight $p_{ij} = 1$ except the edge from node 3 to node 4 which has weight $p_{34} = \delta \in (0, 1)$ and the edge from 3 to 5 with weight $p_{35} = 1 - \delta$. The edge $e_{BA} = (5, 1)$ between $A = 1$ (former source) and $B = 5$ (former sink) has already been added.
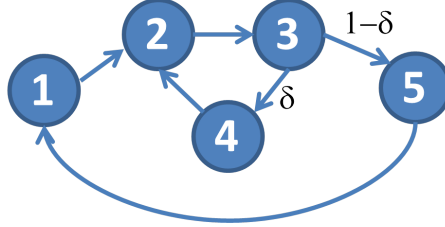

Figure 2: Directed network with $A = 1$, $B = 5$ and edge $e_{BA} = (5, 1)$ already added.

As stated above, we have $q^+ = (0, 1, 1, 1, 1)^T$ and $q^- = (1, 1, 1, 1, 0)^T$. For the transition from A to B and $\delta = 0.25$ we find the following current of reactive trajectories according to (6):

$$F^{AB} = F^+ = \begin{pmatrix} 0 & 0.2000 & 0 & 0 & 0 \\ 0 & 0 & 0.2667 & 0 & 0 \\ 0 & 0 & 0 & 0.0667 & 0.2000 \\ 0 & 0.0667 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and the transition rate $k_{AB} = 0.2$. There are two simple cycles, $\gamma_1 = (2, 3, 4)$ and $\gamma_2 = (1, 2, 3, 5)$. We find $\Gamma_P = \{\gamma_2\}$ and $\Gamma_U = \{\gamma_1\}$ and compute from (4):

$$w(\gamma_1) = 2/30 = 0.6667, \qquad w(\gamma_2) = 0.2,$$

which agrees with (13) since $F = F_{AB}$ everywhere except $F(5, 1) = 0.2 = k_{AB}$.

9

Based on this we can also compute $w(\gamma_2) = k_{AB}$, and $w(\gamma_1)$ as functions of δ, as shown in Fig. 3. The result coincides with simple intuition: The larger δ becomes, the smaller the transition rate k_{AB} gets while simultaneously unproductive flow along cycle γ_1 increases. For $\delta \to 0$ we find $k_{AB} = 0.25$ (because then every transition needs four steps), and $w(\gamma_1) = 0$.
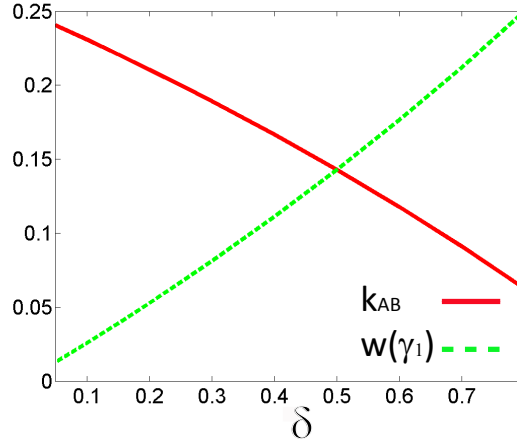

Figure 3: k_{AB} (solid red), $w(\gamma_1)$ (dashed green) as functions of δ.

4.2 The general case

Now we turn to the general case. So let the vertex set $\mathbb{X}$ be decomposed into A, B and T (mutually disjoint), but A and B do not have to be topological sinks/sources any longer and contain more then one vertex. We want to model information transport from A to B again, so A will be the information source in the sense that every time (X_t) visits A, it picks up one bit of information if it does not already have one, and every time it visits B it delivers the bit it carries, if any. This is the typical TPT context, and the number of bits transferred per unit time is again k_{AB}. Now there are two ways in which a trajectory can be unproductive: It can self-intersect in T as before, and it can make many unproductive returns to A before making a transition to B. TPT takes care of the latter: We have a flow decomposition $F = F^{AA} + F^{AB} + F^{BA} + F^{BB}$ where

$$
\begin{aligned}
F_{ij}^{AA} &= q_i^- \mu_i L_{ij}(1 - q_j^+), \\
F_{ij}^{AB} &= q_i^- \mu_i L_{ij} q_j^+, \\
F_{ij}^{BA} &= (1 - q_i^-)\mu_i L_{ij}(1 - q_j^+), \\
F_{ij}^{BB} &= (1 - q_i^-)\mu_i L_{ij} q_j^+.
\end{aligned}
$$

The reactive trajectories from A to B have F^{AB} as their flow. E. Vanden-Eijnden and M. Cameron have shown in [3] that one can write down a generator L^R which generates the reactive trajectories directly and has F^{AB} as its probability flow, meaning $F_{ij}^{AB} = \mu_i^R L_{ij}^R$ where μ^R given by (10) is the invariant distribution of L^R. This is

done by mapping A to a single state a, B to a single state b and then setting

$$L_{ij}^R = L_{ij} \frac{q_j^+}{q_i^+}, \qquad\qquad i,j \in T$$

$$L_{ib}^R = \sum_{j \in B} L_{ij}/q_i^+, \qquad\qquad i \in T \qquad (14)$$

$$L_{aj}^R = \sum_{i \in A} \mu_i L_{ij} q_j^+ /(1 - \rho_R), \qquad\qquad j \in T$$

$$(15)$$

where $\rho_R = \sum_i q_i^- \mu_i q_i^+$. We also add a single edge $e_{ba} = (b,a)$ that reroutes the trajectory to a once it reaches b, and thus set $L_{ba}^R = 1$. That trajectories generated by L^R indeed have F_{ij}^{AB} as their flow can be seen in the following way: By using (10), we get for $i,j \in T$:

$$\mu_i^R L_{ij}^R = q_i^- \mu_i q_i^+ L_{ij} \frac{q_j^+}{q_i^+} = q_i^- \mu_i L_{ij} q_j^+ = F_{ij}^{AB}.$$

The other cases are checked equivalently. By using L^R instead of L, we have reduced the general situation to the simpler one discussed before. The transition matrix corresponding to L^R is $P^R = L^R + I$.

Now the reactive trajectories themselves still have productive and unproductive parts since, as mentioned before, they can still self-intersect in T. In [3] the committor is used to distinguish between productive and unproductive parts. That only works for the reversible case[1]. Instead, we will do the very same construction as before: Any reactive trajectory $c = (a, i_1, \ldots, i_m, a, b)$ can be decomposed into simple loops, exactly one of which contains the edge e_{ba}, and this one we call the productive part. With the splitting of Γ into productive cycles $\Gamma_P = \{\gamma \in \Gamma : e_{BA} \text{ is an edge of } \gamma\}$ and unproductive cycles $\Gamma_U = \{\gamma \in \Gamma : e_{BA} \text{ is not an edge of } \gamma\}$, we can write the decomposition (3) for the flow F^{AB} of reactive trajectories as

$$F^{AB} = \sum_{\gamma \in \Gamma_P} w^R(\gamma) C_\gamma + \sum_{\gamma \in \Gamma_U} w^R(\gamma) C_\gamma = F_P + F_U \qquad (16)$$

where the weights w^R are given in terms of the matrix elements of P^R by

$$w^R(\gamma) = p_{i_1 i_2}^R \ldots p_{i_{s-1} i_s}^R p_{i_s i_1}^R \frac{D^R(\{i_1, \ldots, i_s\})}{\sum_{i \in \mathbb{X}} D^R(\{i\})} \qquad (17)$$

for a cycle $\gamma = (i_1, \ldots, i_s)$. In practical situations, we will use L^R (or P^R) to generate reactive trajectories and sample the weights $w^R(\gamma)$ according to (5). We again have

$$k_{AB} = \sum_{\gamma \in \Gamma_P} w^R(\gamma).$$

---

[1]The reason is that in the reversible case, the effective current (9) can be written as $F_{ij}^+ = \mu_i L_{ij}(q_j^+ - q_i^+)$, so that F_{ij}^+ only flows from lower to higher committor values, and the productive pieces of reactive trajectories can be identified as pieces along which the committor increases. In the non-reversible case, this is no longer true: The committor q^+ may stay constant or even decrease along the flow F^+. In the simple case discussed in section 4.1, $q^+ = 1$ along all reactive trajectories, so q^+ has no information about productive and unproductive parts.

Reversibility and unproductive cycles. If the underlying Markov chain is reversible, then we have $w^R(\gamma) = w^R(\gamma^-)$ for any $\gamma \in \Gamma_U$, where γ^- is γ with reversed orientation[2]. This can be seen in the following way: Let $\gamma = (i_1, \ldots, i_s) \in \Gamma_U$. Then $i_1, \ldots, i_s \in T$, and we have $p_{i_1 i_2} \ldots p_{i_{s-1} i_s} p_{i_s i_1} = p_{i_s i_{s-1}} \ldots p_{i_2 i_1} p_{i_1 i_s}$ from the Kolmogorov criterion for reversibility. From this, we get

$$
\begin{aligned}
p_{i_1 i_2}^R \ldots p_{i_{s-1} i_s}^R p_{i_s i_1}^R &= p_{i_1 i_2} \frac{q_{i_2}^+}{q_{i_1}^+} \ldots p_{i_{s-1} i_s} \frac{q_{i_s}^+}{q_{i_{s-1}}^+} p_{i_s i_1} \frac{q_{i_1}^+}{q_{i_s}^+} \\
&= p_{i_1 i_2} \ldots p_{i_{s-1} i_s} p_{i_s i_1} \\
&= p_{i_s i_{s-1}} \ldots p_{i_2 i_1} p_{i_1 i_s} \\
&= p_{i_s i_{s-1}}^R \ldots p_{i_2 i_1}^R p_{i_1 i_s}^R
\end{aligned}
$$

and finally $w^R(\gamma) = w^R(\gamma^-)$ since the additional factor in (17) does not depend on the ordering of $i_1, \ldots, i_s$. Note that this only holds for unproductive cycles: For $\gamma \in \Gamma_P$, we have $w^R(\gamma^-) = 0$ because γ necessarily contains the edge (ba), and transitions along the reversed edge (ab) are forbidden.

As a consequence, we can see that $F_U = F_U^T$ if the chain is reversible: Since we can organise the cycles in Γ_U into pairs $\{\gamma, \gamma^-\}$ and $C_{\gamma^-} = C_\gamma^T$, we have

$$
F_U = \sum_{\gamma \in \Gamma_U} w(\gamma) C_\gamma = \frac{1}{2} \sum_{\gamma \in \Gamma_U} w^R(\gamma) \left(C_\gamma + C_{\gamma^-} \right) = \frac{1}{2} \sum_{\gamma \in \Gamma_U} w^R(\gamma) \left(C_\gamma + C_\gamma^T \right).
$$

Since F_P and F^{AB} only differ by F^U and F^U is symmetric, the effective current associated to F^{AB} by (9) equals the effective current associated to F^P. Thus in the reversible case our splitting into F_P and F_U does not change the effective current F^+. This is reasonable since in the reversible case, the effective current is already cycle-free and therefore has no unproductive parts.

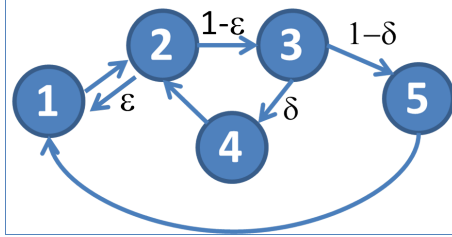**Example, continued.** Let us return to our illustrative example, see Fig. 4.

Figure 4: Directed network with additional flow back to $A = 1$.

For the transition from $A = \{1\}$ to $B = \{5\}$ and $\delta = 0.25$, $\varepsilon = 0.1$ we find the following reactive flow:

$$
F^{AB} = F^+ = \begin{pmatrix}
0 & 0.1888 & 0 & 0 & 0 \\
0 & 0 & 0.2436 & 0 & 0 \\
0 & 0 & 0 & 0.0548 & 0.1888 \\
0 & 0.0548 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix},
$$

---

[2]I.e. if $\gamma = (i_1, \ldots, i_s)$, then $\gamma^- = (i_s, \ldots, i_1)$.

and the transition rate $k_{AB} = 0.1888$. There are three simple cycles $\gamma_1 = (2,3,4)$, $\gamma_2 = (1,2,3,5)$, and $\gamma_3 = (1,2)$; their weight can be computed from (4):

$$w(\gamma_1) = 0.0629, \qquad w(\gamma_2) = 0.1888 = k_{AB}, \qquad w(\gamma_3) = 0.0280.$$

The complete probability flow comes out to be

$$F = \begin{pmatrix} 0 & 0.2186 & 0 & 0 & 0 \\ 0.0280 & 0 & 0.2517 & 0 & 0 \\ 0 & 0 & 0 & 0.0629 & 0.1888 \\ 0 & 0.0629 & 0 & 0 & 0 \\ 0.1888 & 0 & 0 & 0 & 0 \end{pmatrix},$$

with decomposition $F = F^{AA} + F^{AB} + F^{BA} + F^{BB}$ with $F^{BB} = 0$, and

$$F^{AA} = \begin{pmatrix} 0 & 0.0280 & 0 & 0 & 0 \\ 0.0280 & 0 & 0.0081 & 0 & 0 \\ 0 & 0 & 0 & 0.0081 & 0 \\ 0 & 0.0081 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$F^{BA} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.1888 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

On the other hand, the weights w^R generated from L_R in (14) come out to be $w^R(\gamma_1) = 0.0548 < w(\gamma_1)$, $w^R(\gamma_2) = w(\gamma_2) = k_{AB}$ and $w^R(\gamma_3) = 0$. In light of the decomposition (16) we have $\Gamma_P = \{\gamma_2\}$ and $\Gamma_U = \{\gamma_1, \gamma_3\}$. The relation $w^R(\gamma_1) < w(\gamma_1)$ appears since γ_1 also contributes to F^{AA} due to excursions of the form $(1,2,3,4,2,1)$. γ_3 only contributes to F^{AA}, hence $w^R(\gamma_3) = 0$.

Fig. 5 displays the dependence of the reactive, productive flow k_{AB} and of the total unproductive flows $w(\gamma_1)$ and $w(\gamma_3)$ on δ for fixed $\varepsilon = 0.1$. The results are as expected: the flow along γ_3 is constant wrt changes in δ, while k_{AB} decreases and $w(\gamma_1)$ increases with increasing δ.

5 Numerical Examples

5.1 Random walker in a maze

For further illustration we will consider the random walk in the two mazes, a simple one shown in the left panel of Fig. 6 and a slightly more complicated one shown in Fig. 8. In both cases the random walker starts in $A = 1$ and tries to find the exit in $B = 16$. It cannot pass through the walls (thick black lines) and can pass through one-way doors (gray lines) solely in the direction indicated by the blue arrow. If the walker is in a cell it uses every possible exit from the cell with equal probability.

The panel on the right hand side of Fig. 6 shows the probability flow induced by the random walker on the simple maze. By the above construction the flow is decomposed into the reactive flow along two reactive pathways (left hand side panel of Fig. 7) that both have weight 0.02. Consequently the transition rate is $k_{AB} = 0.04$.
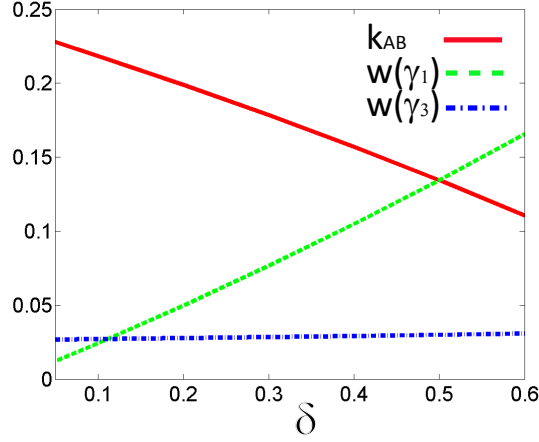
13

Figure 5: k_{AB} (solid red), $w(\gamma_1)$ (dashed green), and $w(\gamma_3)$ (dashed-dotted blue) as functions of δ for $\varepsilon = 0.1$
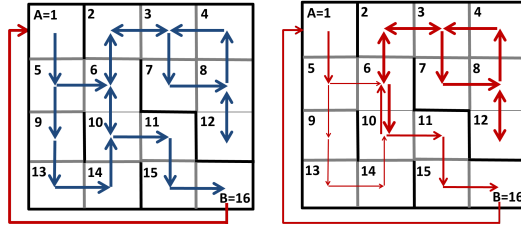

Figure 6: Left: Simple maze with 16 cells and possible moves of the random walker indicated by arrows (cell numbers in black, walls in black, one-way doors in gray). The additional red arrow from B to A indicates the edges e_{BA} along which the random walker is inserted into the maze again after having found the exit. Right: Probability flow induced by the random walker with arrow thickness relative to the flow between cells (thin arrow=0.02, medium=0.04, thick=0.06).

The right hand side panel of Fig. 7 shows the four simple cycles of the decomposition. The cycles $(2, 3)$, $(8, 12)$ and $(3, 7, 8, 4)$ all have weight 0.06, while $(6, 10)$ has weight 0.04.

With this further insight into flow decomposition we can turn to the more complicated maze shown in Fig. 8. This random walk induces the invariant measure μ shown in Fig. 9 together with the forward committor for the transition $A \to B$.

When computing its transition rate by means of TPT we find $k_{AB} = 0.0369$, i.e., the typical transition time is about 27 walker steps. This already shows that the walker runs in cycles somewhat before being able to find the exit. Fig. 10 shows the three reactive pathways with highest weights (together with the edge $e_{BA} = (B, A)$ these form the most important cycles in Γ_P). Together these three contribute more than 50% of the overall transition rate. We observe that the pathways seem to go from lower forward committor to higher forward committor.

When we turn to the unproductive cycles, we observe that the ones carrying the most flow are the "trivial" cycles $(1, 2)$, $(2, 3)$, and $(1, 5)$ with weights $w_1 = 0.049$,
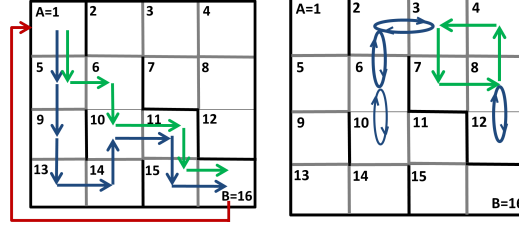
Figure 7: Left: Reactive pathways of random walker in the simple maze; both pathways (blue,green) carry the same weight (0.02). Right: Simple cycles of the simple maze; arrow thickness relative to weight of the respective cycle (thin arrow=0.04 (cycle $(10, 6)$), thick=0.06).
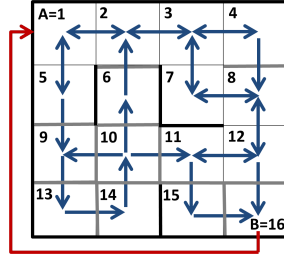

Figure 8: The maze with 16 cells (cell numbers in black, walls in black, one-way doors in gray). Blue arrows indicate the possible directions of exit from a cell. The entrance to the maze is $A = 1$, its exit $B = 16$. The additional red arrow from B to A indicates the edges e_{BA} along which the random walker is inserted back into the maze again after having found the exit.
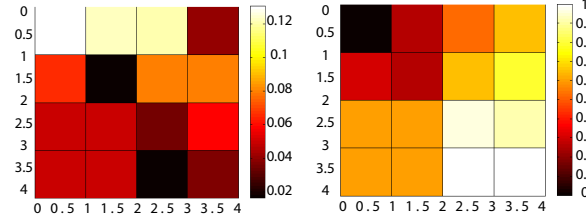

Figure 9: Invariant measure μ (left) and forward committor q^+ of the random walk in the maze.

$w_2 = 0.039$ and $w_3 = 0.033$, respectively. Observe that $(w_1 + w_2 + w_3)/k_{AB} \approx 3.3$, i.e. the three most important unproductive cycles carry about three times more probability flow than all reactive pathways together. The dominant feature of the maze is being trapped in unproductive cycles for long times before successfully finding the exit. The most important nontrivial unproductive cycles and the associated weights are shown in Fig. 11. We observe that they can be combined with the reactive pathways into path-loop structures, i.e., typical reactive trajectories will have parts that progress along a reactive path and other parts that can be seen as loop excursions from the reactive path along unproductive cycles.
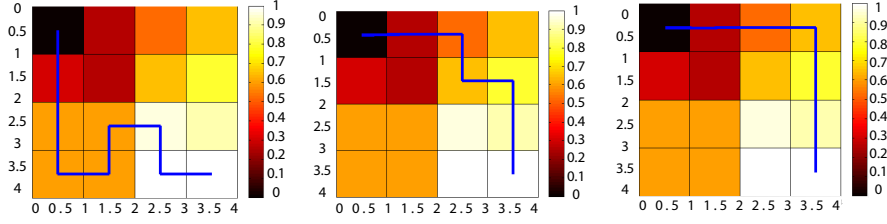
15

Figure 10: First three most important reactive pathways that carry about 22%, 16% and 15% (from left to right) of the overall transition rate. Cell coloring: forward committor.
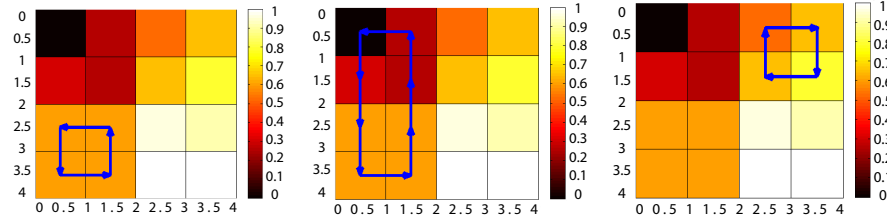

Figure 11: First three most important, nontrivial unproductive cycles, $\gamma_1, \gamma_2, \gamma_3$ with weights $w(\gamma_1) = 0.017$, $w(\gamma_2) = 0.012$, and $w(\gamma_3) = 0.008$. Cell coloring: forward committor.

5.2 Langevin dynamics

Recent years have seen a continuous increase in research on so-called Markov State Models (MSM) for molecular dynamics (MD). MSMs are based on a discretization of the transfer operator associated with MD and aim at the reproduction of the longest timescales in MD, see [25, 2]. However, most of the approaches to MSM building assume that the underlying dynamics is reversible [2], and MSM building for non-reversible MD is still in its infancies.

The most popular model for non-reversible MD is the so-called Langevin System [22]:

$$\dot{r} = p, \qquad \dot{p} = -\nabla_r V(r) - \gamma p + \sigma \dot{W}_t, \tag{18}$$

where r is the vector of the Euclidean coordinates of all atoms in the molecular system, p the vector of associated momenta, $V(r)$ the interaction energy, and $-\nabla_r V(r)$ the vector of all inter-atom interaction forces. $\gamma > 0$ denotes some friction constant and $F_{\text{ext}} = \sigma \dot{W}_t$ the external forcing given by a $3N$-dimensional Brownian motion W_t. The external stochastic force is assumed to model the influence of the heat bath surrounding the molecular system. The total internal energy given by the Hamiltonian $H(r, p) = p^2/2 + V(r)$ is not preserved, but the interplay between stochastic excitation and damping balances the internal energy. As a consequence, the canonical probability density function $\mu(r, p) \propto \exp(-\beta H(r, p))$ with $r = (r, p)$ is invariant w.r.t. the Markov process corresponding to the Langevin system under weak growth conditions on the potential energy function [13], if the noise and damping constants satisfy [22]:

$$\beta = \frac{2\gamma}{\sigma^2}. \tag{19}$$

The Langevin process (Y_t) is not reversible but satisfies the so-called extended detailed balance condition

$$\mathbb{P}\Big(Y_{t+T} = (r', p') \mid Y_t = (r, p)\Big)\, \mu(r, p)$$

$$= \mathbb{P}\Big(Y_{t+T} = (r, -p) \mid Y_t = (r', -p')\Big)\, \mu(r', -p'), \qquad (20)$$

for arbitrary t and $T > 0$.

We consider Langevin dynamics with one-dimensional r and p and double well potential $V(r) = (r^2 - 1)^2$. Figure 12 shows the phase space of this system together with some periodic orbits of the associated Hamiltonian system. In this case the process has two metastable sets around the minima $x_\pm = \pm 1$ of the potential V and $p = 0$. We set $\beta = 5$ and $\gamma = 0.2$, so that the Langevin dynamics is still "close" to the associated Hamiltonian system

$$\dot{r} = p, \qquad \dot{p} = -\nabla_r V(r), \qquad (21)$$

i.e., if we start the Langevin process in (r_0, p_0) with energy $E_0 = H(r_0, p_0) < 0.9$ then the dynamics will approximately follow the periodic orbit $H(r, p) = E_0$ (with $r < 0$ if $r_0 < 0$ and $r > 0$ if $r_0 > 0$) of the associated Hamiltonian system for some time interval of order 1. Therefore the typical transition from the vicinity of one of the wells across the energy barrier at $r = 0$ towards the other well will look as follows: First the trajectory will orbit the initial well for some period of time before it crosses the barrier and starts to orbit the target well until it finally hits the close vicinity around the respective energy minimum.
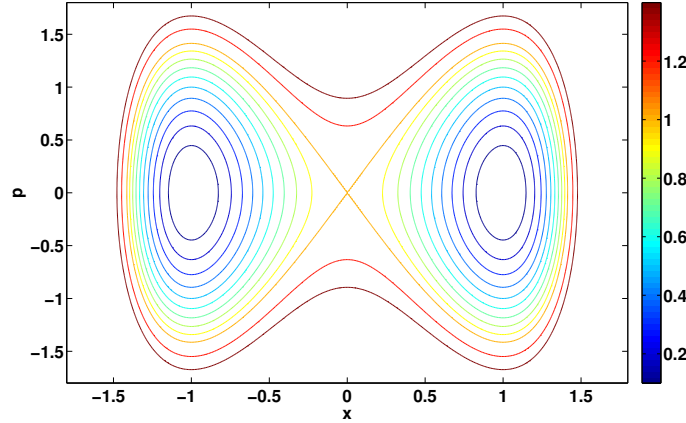

Figure 12: Some periodic orbits of the system with Hamiltonian $H = \frac{1}{2}p^2 + V(r)$. The coloring is according to the energy $E_0 = H(r_0, p_0)$.

For this Langevin process MSM building is done as follows [2], [25]: We constructed a uniform box covering of the essential state space $[-1.8, 1.8] \times [-1.8, 1.8]$ where the invariant pdf is larger than the square root of the machine precision. We took square boxes $B_i, i = 1, \ldots, n$ of size $\Delta r = 0.2$ and $\Delta p = 0.2$. Next, $M = 100$ trajectories of length $\tau = 0.25$ were started in every box with μ-distributed initial points, and the transition matrix

$$P_{ij} = \frac{m_{ij}}{M}$$

17

was constructed, where m_{ij} is the number of trajectories starting in B_i and ending up in B_j. The length $\tau = 0.25$ of the trajectories is very small compared to the expected transition time (which is larger than 100 here) and still shorter than the period of the periodic orbits of the associated Hamiltonian system. MSM [25] theory tells us that the leading eigenvalues of the transition matrix P are very close approximations of the leading eigenvalues of the Langevin transfer operator and thus allow for an approximation of the transition statistics between the metastable wells.

We define the two core sets $A = \{0.9 \leq r \leq 1.1, -0.3 \leq p \leq 0.3\}$ and $B = \{-1.1 \leq r \leq -0.9, -0.3 \leq p \leq 0.3\}$ and consider the current F^{AB}, see (6), of reactive trajectories from A to B (here $L = P - I$). The decomposition (16) should result in many unproductive cycles that approximately correspond to orbits of the Hamiltonian system around $y_{\pm} = (r_{\pm}, 0)$. To demonstrate this, we sampled from the ensemble of reactive trajectories. We have $k_{AB} = 0.0014$, hence we are able to observe ≈ 1.400 transitions with a realization of length $T = 10^6$. The cycle decomposition (16) is then computed numerically.
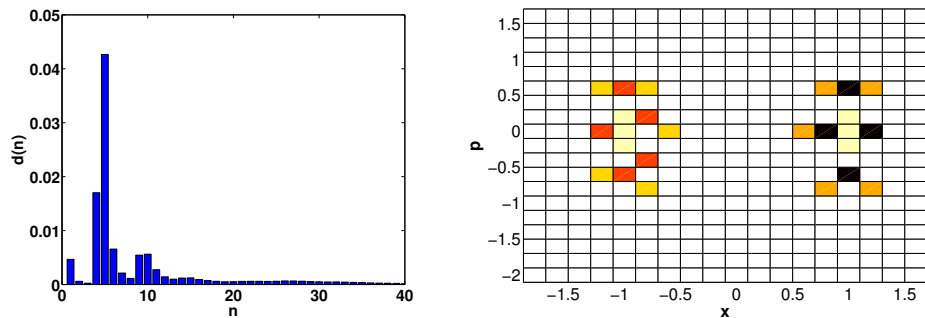

Figure 13: Left: Empirical distribution $d(n)$ of the lenghts of cycles in Γ_U. Right: The core sets A and B (light yellow) and γ_1, γ_4, γ_{13} and γ_{15}.

To obtain a first insight into the structure of Γ_U, we compute the empirical distribution of cycle lengths $d(n) = \sum_{\gamma \in \Gamma_U} w(\gamma)\delta(|\gamma| = n)$. The number $T \cdot d(n)$ is the number of times an unproductive cycle of length n was passed through by the realization. The result is shown in Figure 13. We observe that Γ_U is strongly dominated by cycles of length 4 and 5. Next we sort cycles in Γ_U according to $w(\gamma)$, let $\gamma_i \in \Gamma_U$ be the cycle with ith largest weight. In Figure 13 on the right, we show $\gamma_1, \gamma_4, \gamma_{13}$ and γ_{15} together with A and B to illustrate that these cycles track orbits that wind once around A and B. All other cycles γ_i with $1 \leq i \leq 20$ look similar, but overlap partially with the ones shown. These results are to be expected: The Langevin system is close to the Hamiltonian system which completes orbits with winding number one in time $t = 1$, since the MSM time step is $\tau = 0.25$ this should correspond to cycles of length 4 or 5.

Next we consider the probability current of reactive trajectories F^{AB} and its decomposition $F^{AB} = F_P + F_U$. In Figure 14, the effective current F^+, see (9), and the productive current F_P are shown. We observe that F^+ is strongly dominated by rotations around A and B. The productive current F_P has less pronounced rotational parts and shows the emergence of a transition channel in the $p < 0$ half plane.
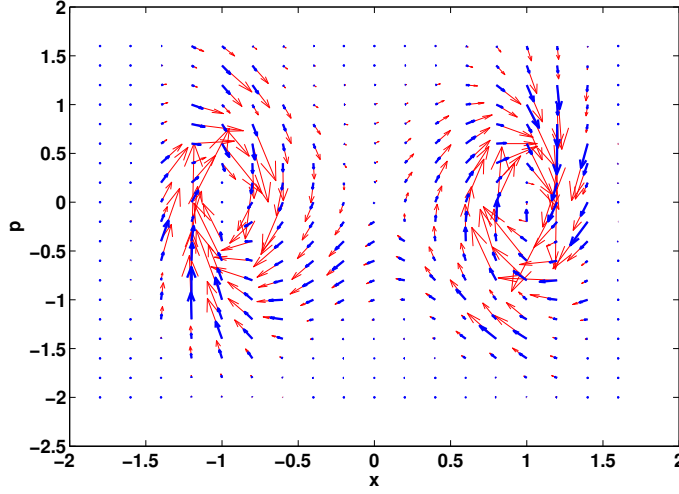
Figure 14: Effective current F^+ (red) and productive current F_P (blue) of the Langevin system.

6   Conclusion

We have presented a comprehensive theory for solving the problem of characterizing the reactive/productive and the unproductive flow in transitions between an initial and a target set for undirected and directed networks and associated Markov processes. The theory can be seen as an extension of transition path theory (TPT) that uses the stochastic cycle decomposition of a network for distinguishing between reactive pathways (productive flow) and additional cycles that carry part of the reactive flow but contribute to the transition in an unproductive way only.

Our approach allows to uncover the structure of a reactive event in more detail than TPT. The additional information about cycles in which parts of the reactive flow runs in loops is essential, e.g., when one is interested in influencing the reactive event in a certain way (improving transition rates or avoiding metastable traps). The last example, in which we use the new approach for analyzing Langevin dynamics, demonstrates this aspect in a nutshell: the information on the cycles allows to understand exactly to which extent the periodic orbits of the underlying Hamiltonian system still contribute to the transition process (for small enough friction and noise).

Finally, we mention some open problems:

- If the system considered has a large number of states, then there will be very many different pathways from A to B that each contribute very little to the overall reaction. In such a situation, information about individual pathways is of little use unless we have a way of identifying groups of pathways which are similar.

- While our theory gives a rather complete picture of the underlying structure of transitions between initial and target sets in networks and Markov chains, the efficiency of the algorithms presented here does not allow an application to really large networks: The determination of the cycle weights requires sam-

19

pling of transition events which is inefficient for strongly metastable processes. Thus, improving the efficiency of the algorithms presented here needs further investigation.

References

[1] B. Altaner, S. Grosskinsky, S. Herminghaus, L. Katthän, M. Timme, and J. Vollmer. Network representations of nonequilibrium steady states: Cycle decompositions, symmetries, and dominant paths. *Phys. Rev. E*, 85:041133, 2012.

[2] G. R. Bowman, V. S. Pande, and F. Noé, editors. *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, volume 797 of *Advances in Experimental Medicine and Biology*. Springer, 2014.

[3] M. Cameron and E. Vanden-Eijnden. Flows in complex networks: Theory, algorithms, and application to Lennard-Jones cluster rearrangement. *Journal of Statistical Physics*, 156(3):427–454, 2014.

[4] J. Chodera, N. Singhal, V. S. Pande, K. Dill, and W. Swope. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *Journal of Chemical Physics*, 126, 2007.

[5] N. Djurdjevac-Conrad, R. Banisch, and Ch. Schuette. Modularity of directed networks: Cycle decomposition approach. *submitted to J. Comp. Dynamics*, 2014. arXiv:1407.8039v2 [math-ph].

[6] W. E and E. Vanden-Eijnden. Metastability, conformation dynamics, and transition pathways in complex systems. In *Multiscale modelling and simulation*, volume 39 of *Lect. Notes Comput. Sci. Eng.*, pages 35–68. Springer, Berlin, 2004.

[7] Weinan E and E. Vanden-Eijnden. Towards a theory of transition paths. *Journal of statistical physics*, 123:503–523, 2006.

[8] Weinan E and E. Vanden-Eijnden. Transition-path theory and path-finding algorithms for the study of rare events. *Annual Review of Physical Chemistry*, 61:391–420, 2010.

[9] D. Jiang, M. Qian, and M.-P. Quian. *Mathematical theory of nonequilibrium steady states: on the frontier of probability and dynamical systems.* Springer, 2004.

[10] Sophia L. Kalpazidou. *Cycle Representations of Markov Processes.* Springer, 2006.

[11] Kai J Kohlhoff, Diwakar Shukla, Morgan Lawrenz, Gregory R Bowman, David E Konerding, Dan Belov, Russ B Altman, and Vijay S Pande. Cloud-based simulations on google exacycle reveal ligand modulation of gpcr activation pathways. *Nature chemistry*, 6(1):15–21, 2014.

[12] Tijun Li, Weinan E, and Eric Vanden Eijnden. Optimal partition and effective dynamics of complex networks. *Proc. Nat. Acad. Sci.*, 105, 2008.

[13] J. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerated noise. *Stochastic Process Appl.*, 101 (2):185–232, 2002.

[14] P. Metzner. *Transition Path Theory for Markov Processes: Application to molecular dynamics.* PhD thesis, Free University Berlin, 2007.

[15] P. Metzner, Ch. Schütte, and E. Vanden-Eijnden. Illustration of transition path theory on a collection of simple examples. *J. Chem. Phys.*, 125(8), 2006. 084110.

[16] P. Metzner, Ch. Schütte, and E. Vanden-Eijnden. Transition path theory for Markov jump processes. *Multiscale Modeling and Simulation*, 7(3):1192–1219, 2009.

[17] F. Noe and S. Fischer. Transition networks for modeling the kinetics of conformational change in macromolecules. *Curr. Opin. Struct. Biol.*, 18:154–162, 2008.

[18] F. Noe, I. Horenko, C. Schuette, and J.C. Smith. Hierarchical analysis of conformational dynamics in biomolecules: Transition networks of metastab le states. *J. Chem. Phys.*, 126:155102, 2007.

[19] F. Noé, Ch. Schütte, E. Vanden-Eijnden, L. Reich, and T. Weikl. Constructing the full ensemble of folding pathways from short off-equilibrium trajectories. *PNAS*, 106(45):19011–19016, 2009.

[20] Albert C. Pan and Benoit Roux. Building Markov state models along pathways to determine free energies and rates of transitions. *J Chem Phys.*, 129(6):064107, 2008.

[21] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Fischbach, M. Held, J. D. Chodera, Ch. Schütte, and F. Noé. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.*, 134:174105, 2011.

[22] H. Risken. *The Fokker-Planck Equation.* Springer, New York, 1996. 2nd edition.

[23] M. Sarich, N. Djurdjevac, S. Bruckner, T. O. F. Conrad, and Ch. Schütte. Modularity revisited: A novel dynamics-based concept for decomposing complex networks. *Journal of Computational Dynamics*, 1(1):191–212, 2014.

[24] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.*, 48:571–585, Oct 1976.

[25] Ch. Schütte and M. Sarich. *Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches*, volume 24 of *Courant Lecture Notes.* American Mathematical Society, December 2013.

[26] C. R. Schwantes, R. T. McGibbon, and V. S. Pande. Perspective: Markov models for long-timescale biomolecular dynamics. *The Journal of Chemical Physics*, 141(9):–, 2014.

[27] E. Vanden-Eijnden. Transition path theory. In Mauro Ferrario, Giovanni Ciccotti, and Kurt Binder, editors, *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, volume 703 of *Lecture Notes in Physics*, pages 453–493. Springer Berlin Heidelberg, 2006.

[28] H. Wang and Ch. Schütte. Building markov state models for periodically driven non-equilibrium systems. *accepted for Journal of Chemical Theory and Computation*, 2015.

[29] R. K. P. Zia and B. Schmittmann. Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states. *Journal of Statistical Mechanics-theory and Experiment*, 7, 2007.