Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

M. SARICH[1], CHRISTOF SCHÜTTE[1,2]

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

[2] *Zuse Institute Berlin, Germany*

Utilizing hitting times for finding metastable sets in non-reversible Markov chains

ZIB-Report 14-32 (August 2014)

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7
D-14195 Berlin-Dahlem

Telefon: 030-84185-0
Telefax: 030-84185-125

e-mail: bibliothek@zib.de
URL: http://www.zib.de

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Utilizing hitting times for finding metastable sets in non-reversible Markov chains

M. Sarich[1,3] Ch. Schuette[1,2]

August 1, 2014

Abstract

Techniques for finding metastable or almost invariant sets have been investigated, e.g., for deterministic dynamical systems in set-oriented numerics, for stochastic processes in molecular dynamics, and for random walks on complex networks. Most prominent algorithms are based on spectral apporaches and identify metastable sets via the doimant eigenvalues of the transfer operator associated with the dynamical system under consideration. These algorithms require the dominant eigenvalues to be real-valued. However, for many types of dynamics, e.g. for non-reversible Markov chains, this condition is not met. In this paper we utilize the hitting time apporach to metastable sets and demonstrate how the well-known statements about optimal metastable decompositions of reversible chains can be reformulated for non-reversible chains if one switches from a spectral approach to an exit time approach. The performance of the resulting algorithm is illustrated by numerical experiments on random walks on complex networks.

Keywords: metastability, hitting times, non-reversible Markov chain, directed networks, random walk

Mathematical Subject Classification: 60J22 / 60J20, Secondary: 37M99

1 Introduction

In recent years techniques for finding metastable or almost invariant sets in dynamical systems have attracted a lot of attention, for deterministic systems [6] and set-oriented numerics [5, 7] as well as stochastic systems [18] with applications in molecular dynamics [19, 3]. These techniques allow to identify metastability by discretization of the transfer operator or Frobenius Perron operator of the underlying dynamics which results in the a stochastix matrix and an associated Markov chain. Metastable or almost invariant sets are then identified via spectral approaches: the dominant eigenvectors of the matrix contain essential information about the invariant measure of the system and about the decomposiition of the system into metastable sets.

[1]BioComputing Group, Institute of Mathematics, Freie Universität Berlin
[2]Zuse Institute Berlin (ZIB), Takustr. 7, 14195 Berlin
[3]Corresponding author. Email: sarich@math.fu-berlin.de

There is a long list of articles on finding metastable decompositions of Markov chains, via spectral approaches [9, 12, 10, 11, 8, 13, 19] and Markov chain aggregation techniques [4, 15] as well as via exit time approaches [1, 2, 19]. However, by far most results are only available for reversible Markov chains, or do not result in robust algorithms for finding metastable decompositions with more than two sets.

In addition to the discussion about metastability in dynamical systems Markov chain decompositions are key to finding good partitions of networks into modules: When wanting to find the strongly connected modules of a network, one can do this via the metastable sets of an appropriately defined random walk on the network, see [14, 17]. The Markov chain assoiated with the random walk is reversible as long as the network is undirected; directed network lead to non-reversible chains. Therefore many of the powerful algorithms for identifying metastability cannot be used for finding modules in directed networks.

This situation calls for a generalization of the theory for decomposing metastable Markov chains to non-reversible chains. This article presents such a generalization. It is shown that the well-known statement about optimal metastable decompositions of reversible chains can be reformulated for non-reversible chains if one switches from a spectral approach to an exit time approach: We can get a lower bound to the metastability index of a decomposition by considering the exit time distributions from metastable subsets such that decompositions with strong metastability can be found by lower bound maximization. Furthermore, we will see that the distribution of hitting times of test sets are almost constant on any metastable set with sharp jumps between sets; consequently we can use this property for identifying metastable sets. After establishing the underlying theory we will show how to use the results algorithmically and demonstrate the performance of the resulting algorithms in some numerical examples.

2 Setting

In all of the following we consider an irreducible aperiodic homogeneous Markov chain (X_j) on state space $\mathbb{X} = \{1, \ldots, n\}$ with transition matrix P and transition probability $p(x, y)$ for $x, y \in \mathbb{X}$. We assume that μ denotes the invariant measure of the chain so that $\mu^T = \mu^T P$. The invariant measure defines the μ-weighted scalar product $\langle u, v\rangle_\mu = \sum_{x\in\mathbb{X}} u(x)v(x)\mu(x)$ and the associated weighted 2-norm $\|\cdot\|_\mu$. For any given set $A \subset \mathbb{X}$ we define the probability to stay in A during one step by

$$P(A) = \mathbb{P}_\mu \left[X_1 \in A | X_0 \in A\right],$$

where the index μ indicates that X_0 is distributed accoring to μ. Using the weighted scalar product the probability to stay in A can be written

$$P(A) = \frac{\langle \mathbf{1}_A, P\mathbf{1}_A\rangle_\mu}{\mu(A)}$$

where $\mu(A) = \sum_{x\in\mathbb{X}} \mu(x) = \langle \mathbf{1}_A, \mathbf{1}_A\rangle_\mu$ is the invariant measure of A and $\mathbf{1}_A$ denotes the indicator function of set A.

2.1 Metastability

Intuitively any metastable set should satisfy $P(A) \approx 1$ with metastability the stronger the closer $P(A)$ is to 1. Consequently, the chain has a metastable decomposition $\mathcal{D} = \{A_1, \ldots, A_m\}$ into m disjoint sets with $\cup_i A_i = \mathbb{X}$ if the metastability index $\mathcal{M}(\mathcal{D})$ of the decomposition,

$$\mathcal{M}(\mathcal{D}) = \frac{1}{m}\left(P(A_1) + \ldots + P(A_m)\right),$$

is close to one. Alternatively, if a set is metastable we intuitively also expect that the exit time of the chain from the set is large. In order to analyse this relationship let us first define the hitting time of a set $B \subset \mathbb{X}$ by $\tau(B) = \min\{k : X_k \in B\}$, and its expectation value $\mathbb{E}_x(\tau(B)) = \mathbb{E}(\tau(B)|X_0 = x)$ when started in $x \notin B$ wrt the law of (X_j). Then, the expected exit time from A is given by the hitting time $\tau(A^c)$ of the complement $A^c = \mathbb{X} \setminus A$ if started somewhere in A, and the expected exit time is

$$\mathbb{E}_x(\tau(A^c)) = \mathbb{E}(\tau(A^c)|X_0 = x), \qquad x \in A.$$

2.2 Networks

Consider a strongly connected graph/network $G = (V, E)$, where V is the set of n nodes and E the set of edges of the graph. We denote the adjacency matrix of the network by $(a(x,y))_{x,y\in V}$ and the out-degree of a node x by $d(x) = \sum_{y\in\mathbb{X}} a(x,y)$. If the network has weighted edges then $a(x,y)$ denotes the weight of the edge $(x,y) \in E$ and $a(x,y) = 0$ if $(x,y) \notin E$. Whenever the network is undirected the adjacency matrix is symmetric. 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,

$$p(x,y) = \frac{a(x,y)}{d(x)}. \tag{1}$$

The associated Markov chain is irreducible and aperiodic, and if reversible, it has invariant measure $\mu(x) = d(x)/\sum_x d(x)$, i.e., nodes with high degree are visited often, i.e., subnetworks with strong internal connectedness are metastable sets of the associated chain. However, other random walks have been considered also where $p(x,y)$ is defined in terms of $a(\cdot,\cdot)$ in a different way [17].

3 Metastable Decomposition of Reversible Markov Chains

Whenever the Markov chain is reversible, its metastability can be analysed based on the leading eigenvalues of P. The chain (X_j) is said to be reversible if the detailed balance condition $\mu(x)p(x,y) = \mu(y)p(y,x)$ holds. Then, P is symmetric wrt. $\langle\cdot,\cdot\rangle_\mu$ and thus all its eigenvalues are real-valued. The standard random walk for any undirected network is reversible.

For reversible chains, several mathematical statements relating dominant eigenvalues, the corresponding eigenfunctions and a decomposition of the state space into metastable subsets are available [9, 12, 10, 13, 2, 19]. We will here consider the statement that is based on the metastability index $\mathcal{M}(\mathcal{D})$ of a decomposition $\mathcal{D} = \{A_1, \ldots, A_m\}$ of $\mathbb{X}$, see [12], and for a more general version [19]:

Theorem 1 *Let P denote a reversible $n \times n$ transition matrix with lowest eigenvalue $a = \min \sigma(P) > -1$. Let with $\lambda_m \leq \ldots \leq \lambda_2 < \lambda_1 = 1$ be its eigenvalues, possibly counted according to multiplicity. Denote by $v_m, \ldots, v_1$ the corresponding eigenfunctions, normalized to $\|v_k\|_2 = 1$. Let Q be the orthogonal projection of $\mathbb{R}^n$ onto $\mathrm{span}\{\mathbf{1}_{A_1}, \ldots, \mathbf{1}_{A_m}\}$. The metastability of an arbitrary decomposition $\mathcal{D} = \{A_1, \ldots, A_m\}$ of the state space $\mathbb{X}$ can be bounded from above by*

$$P(A_1) + \ldots + P(A_m) \leq 1 + \lambda_2 + \ldots + \lambda_m,$$

while it is bounded from below according to

$$1 + \kappa_2\lambda_2 + \ldots + \kappa_m\lambda_m + c \leq P(A_1) + \ldots + P(A_m)$$

where

$$\kappa_j = \|Qv_j\|_\mu^2 = \sum_{k=1}^{m} \frac{1}{\mu(A_k)} \langle v_j, \mathbf{1}_{A_k} \rangle_\mu^2.$$

and $c = a\,(1 - \kappa_2) \ldots (1 - \kappa_n)$.

Theorem 1 states that the metastability of an arbitrary decomposition $\mathcal{D}$ cannot be larger than $1+\lambda_2+\ldots+\lambda_m$, while it is at least $1+\kappa_2\lambda_2+\ldots+\kappa_m\lambda_m+c$, which is close to the upper bound whenever $\kappa_j \approx 1$ for all $j = 2, \ldots, n$.

Thus the metastability index of a decomposition $A_1, \ldots, A_m$ will be high if

- the eigenvalues $\lambda_2 \geq \ldots \geq \lambda_m$ are all close to 1, and
- the dominant eigenfunctions $v_2, \ldots, v_m$ are almost constant on the metastable subsets $A_1, \ldots, A_m$ implying $\kappa_j \approx 1$.

The term c can be interpreted as a small correction whenever $a \approx 0$ or $\kappa_j \approx 1$. It is demonstrated in [12] that the lower and upper bounds are sharp and asymptotically exact.

Theorem 1 highlights the strong relation between a decomposition of the state space into metastable subsets and dominant eigenvalues close to 1. Studies on the relation between nearly uncoupled Markov chains and dominant eigenvalues yield similar statements, cf. [15, 9, 2, 16].

In view of Theorem 1, it is natural to ask, whether there is an *optimal* decomposition with highest possible metastability. Several algorithms have been proposed to solve the associated optimization problem. However, even if there exists a unique optimal decomposition, the problem of finding it might be ill–conditioned [19]. The reason for this is that there may be extended transition regions between the metastable core sets for which a distinct assignment to one of the metastable core sets does not make sense. This problem can be resolved by relaxing the optimization to finding the core sets *after* identification and extraction of the transition region [19].

4 Metastable Decomposition of General Markov Chains

Whenever we consider a directed network, the associated random walk in general is not reversible. For non-reversible chains the above and comparable results do not allow for finding a metastable decomposition. In order to do so we aim at reformulating the result in terms of exit times instead of spectral elements.

4.1 Exit times from metastable sets

First we aim at a result that connects the probability to stay in a set A to the mean exit time from it.

Lemma 1 *For every set $A \subset \mathbb{X}$ we have that*

$$\frac{\mathbb{E}_A[\tau(A^c)] - 1}{\max\limits_{x \in A} \mathbb{E}_x[\tau(A^c)]} \leq P(A),$$

where

$$\mathbb{E}_A[\tau(A^c)] = \frac{1}{\mu(A)} \sum_{x \in A} \mu(x) \mathbb{E}_x[\tau(A^c)].$$

denotes the μ-averaged exit time from A.

Proof. Define the mean exit time form set A starting in x by $m(x) = \mathbb{E}_x[\tau(A^c)]$. Then for any $x \in A$ it is well-known that

$$m(x) = 1 + \sum_{y \in A} p(x,y) m(y).$$

Inserting this into the definition of $\mathbb{E}_A[\tau(A^c)]$ we have

$$\mathbb{E}_A[\tau(A^c)] = 1 + \frac{1}{\mu(A)} \sum_{x,y \in A} \mu(x) p(x,y) m(y).$$

Dividing by $M := \max\limits_{x \in A} \mathbb{E}_x[\tau(A^c)]$ yields

$$\begin{aligned} \frac{\mathbb{E}_A[\tau(A^c)] - 1}{M} &= \frac{1}{\mu(A)} \sum_{x,y \in A} \mu(x) p(x,y) \underbrace{\frac{m(y)}{M}}_{\leq 1} \\ &\leq \frac{1}{\mu(A)} \sum_{x,y \in A} \mu(x) p(x,y) = P(A). \end{aligned}$$

□

Lemma 1 exhibits that we will have $P(A) \approx 1$ if

- the mean exit time $\mathbb{E}_A[\tau(A^c)]$ from A is large, i.e., $\mathbb{E}_A[\tau(A^c)]^{-1}$ is a small number, and
- $\mathbb{E}_x[\tau(A^c)]$ is almost constant for all $x \in A$ such that $\max_x \mathbb{E}_x[\tau(A^c)] \approx \mathbb{E}_A[\tau(A^c)]$.

As a simple consequence of this lemma, we get the following lower bound to the metastability index of an arbitrary decomposition

Theorem 2 *Let $A_1, ..., A_m$ be a decomposition of $\mathbb{X}$. Then,*

$$\hat{\lambda}_1\hat{\kappa}_1 + \cdots + \hat{\lambda}_n\hat{\kappa}_n \leq \sum_{i=1}^{m} P(A_i)$$

where

$$\hat{\lambda}_i = 1 - \frac{1}{\mathbb{E}_{A_i}[\tau(A_i^c)]}, \qquad \hat{\kappa}_i = \frac{\mathbb{E}_{A_i}[\tau(A_i^c)]}{\max\limits_{x\in A_i} \mathbb{E}_x[\tau(A_i^c)]}.$$

and $\mathbb{E}_{A_i}[\tau(A_i^c)]$ as defined in Lemma 1.

Proof. We define $m_i(x) = \mathbb{E}_x[\tau(A_i^c)]$ and $M_i := \max\limits_{x\in A_i} \mathbb{E}_x[\tau(A_i^c)]$ and find immediately that

$$\frac{\mathbb{E}_{A_i}[\tau(A_i^c)] - 1}{M_i} = \hat{\lambda}_i\hat{\kappa}_i.$$

Using Lemma 1 for every set A_i and summation over $i = 1, ..., m$ completes the proof. □

According to this result, a decomposition $A_1, ..., A_m$ has a high metastability index if

- all $\hat{\lambda}_i$ are close to 1, that is, if the expected exit times of the sets A_i are large on average with respect to the invariant measure, and
- all $\hat{\kappa}_i$ are close to 1, that is, if the averaged expected exit time of each set is close to the largest possible exit time from single points within the respective set.

The latter condition is definitely fulfilled if each function $m_i(x) = \mathbb{E}_x[\tau(A_i^c)]$ of exit times is almost constant on the respective set A_i since

$$\hat{\kappa}_i = \frac{1}{\max\limits_{x\in A_i} m_i(x)} \frac{1}{\mu(A_i)} \sum_{x\in A_i} m_i(x)\mu(x).$$

Note that $\hat{\kappa}_i$ only measures the deviation of the μ-average of expected exit time on the set A_i to its maximal value, but there is no explicit relation to the minimal value. That is, the existence of states with rather small exit time from set A_i do not affect the value of $\hat{\kappa}_i$ as long as the invariant measure on these states is low enough such that they do not strongly contribute to the average exit time.

Summarizing we see that $\hat{\lambda}_i$ and $\hat{\kappa}_i$ take the roles of λ_i and κ_i in Thm. 1.

Example: For reversible chains, the lower bounds of Thms. 1 and 2 can be compared directly. To this end we consider the reversible 4-state chain with transition matrix

$$P = \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.5 & 0.5-\delta & \delta & 0 \\ 0 & \delta & 0.4-\delta & 0.6 \\ 0 & 0 & 0.3 & 0.7 \end{pmatrix},$$

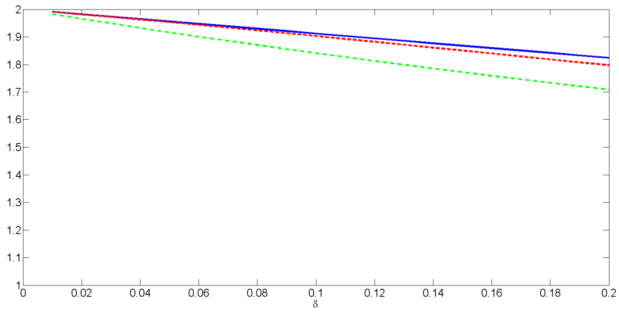

Figure 1: Metastable decomposition of the reversible Markov chain discussed in the text. Solid line: $P(A_1) + P(A_2)$ as a function of δ. Dashed line: Spectral bound $1 + \lambda_2\kappa_2$ as a function of δ. Dashed-dotted line: Hitting time bound $\hat{\lambda}_2\hat{\kappa}_1 + \hat{\lambda}_2\hat{\kappa}_2$ as a function of δ.

where $0 \leq \delta < 0.4$ acts as a free parameter. For small δ the sets $A_1 = \{1, 2\}$ and $A_2 = \{3, 4\}$ form a metastable decomposition. Figure 1 shows the dependence of the metastability index on δ in comparison to the two lower bounds of Thms. 1 and 2.

4.2 Hitting times with respect to test sets

Next we will present conditions under which $\hat{\kappa}_i$ is close to 1, i.e., the function $m_i(x) = \mathbb{E}_x[\tau(A_i^c)]$ of exit times is almost constant on the respective set A_i.

As a preparation we need

Lemma 2 *It holds*

$$\mathbb{E}_x[\tau(D) \mid \tau(D) > \tau(y)] = \mathbb{E}_x[\tau(y) \mid \tau(D) > \tau(y)] + \mathbb{E}_y[\tau(D)].$$

Proof. Let

$$\mathcal{T} = \{(x_0, x_1, ..., x_n) \mid n \in \mathbb{N}, x_n \in D, x_i \notin D, i < n, \exists k < n, x_k = y\}$$

denote the space or all finite trajectoies that lead from any initial point x to D and hit y before D. For $t \in \mathcal{T}$ denote the length of the trajectory by $n(t)$, so

$$n(t) = n - 1 \;\Leftrightarrow\; t = (x_0, x_1, ..., x_n) \in ,x_n \in D, x_i \notin D, i < n.$$

Then,

$$\mathbb{E}_x[\tau(D) \mid \tau(D) > \tau(y)] = \sum_{t \in \mathcal{T}} n(t)\mathbb{P}_x[t].$$

Let

$$\mathcal{T}^- = \{(x_0, x_1, ..., x_n) \mid n \in \mathbb{N}, x_n = y, x_i \neq y \text{ and } x_i \notin D, i < n\}$$

and

$$\mathcal{T}^+ = \{(x_0, x_1, ..., x_n) \mid n \in \mathbb{N}, x_0 = y, x_n \in D, x_i \notin D, i < n\}.$$

Then the concatenation $\circ : \mathcal{T}^- \times \mathcal{T}^+ \to \mathcal{T}$

$$(r_0, r_1, ..., r_k, y) \circ (y, s_1, ..., s_l) = (r_0, r_1, ..., r_k, y, s_1, ..., s_l)$$

is obviously a bijection with $n(t^- \circ t^+) = n(t^-) + n(t^+)$. So,

$$\mathbb{E}_x[\tau(D) \mid \tau(D) > \tau(y)] = \sum_{t^- \in \mathcal{T}^-, t^+ \in \mathcal{T}^+} n(t^- \circ t^+)\mathbb{P}_x[t^- \circ t^+].$$

Now, $\mathbb{P}_x[t^- \circ t^+] = \mathbb{P}_x[t^+|t^-]\mathbb{P}_x[t^-]$. By the Markov property $\mathbb{P}_x[t^+|t^-] = \mathbb{P}_y[t^+]$. That is,

$$\begin{aligned}
\mathbb{E}_x[\tau(D) \mid \tau(D) > \tau(y)] &= \sum_{t^- \in \mathcal{T}^-, t^+ \in \mathcal{T}^+} (n(t^-) + n(t^+))\mathbb{P}_x[t^-]\mathbb{P}_y[t^+] \\
&= \sum_{t^- \in \mathcal{T}^-} n(t^-)\mathbb{P}_x[t^-] \underbrace{\sum_{t^+ \in \mathcal{T}^+} \mathbb{P}_y[t^+]}_{=1} + \sum_{t^+ \in \mathcal{T}^+} n(t^+)\mathbb{P}_y[t^+] \underbrace{\sum_{t^- \in \mathcal{T}^-} \mathbb{P}_x[t^-]}_{=1} \\
&= \mathbb{E}_x[\tau(y) \mid \tau(D) > \tau(y)] + \mathbb{E}_y[\tau(D)].
\end{aligned}$$

□

Now we can prove the main

Theorem 3 *Consider $x, y \in \mathbb{X}$ and a test set $D \subset \mathbb{X}$. Then, the following two inequalities hold:*

(i)

$$\mathbb{E}_x[\tau(D)] \leq \mathbb{E}_x[\tau(y)] + \mathbb{E}_y[\tau(D)]$$

(ii)

$$\frac{|\mathbb{E}_x[\tau(D)] - \mathbb{E}_y[\tau(D)]|}{\max\{\mathbb{E}_x[\tau(D)], \mathbb{E}_y[\tau(D)]\}} \leq \max\left\{\frac{\mathbb{E}_x[\tau(y)]}{\mathbb{E}_x[\tau(D)]}, \frac{\mathbb{E}_y[\tau(x)]}{\mathbb{E}_y[\tau(D)]}\right\}.$$

Proof. Let $q_x(D) = \mathbb{P}_x[\tau(D) < \tau(y)]$ and $q_x(y) = \mathbb{P}_x[\tau(D) > \tau(y)] = 1 - q_x(D)$ denote the committor probabilities for the two sets D and $\{y\}$. Then,

$$\mathbb{E}_x[\tau(D)] = \mathbb{E}_x[\tau(D)|\tau(D) < \tau(y)]q_x(D) + \mathbb{E}_x[\tau(D)|\tau(D) > \tau(Y)]q_x(y).$$

Lemma 2 yields

$$\mathbb{E}_x[\tau(D)] = \mathbb{E}_x[\tau(D)|\tau(D) < \tau(y)]q_x(D) + \mathbb{E}_x[\tau(y) \mid \tau(D) > \tau(y)]q_x(y) + \mathbb{E}_y[\tau(D)]q_x(y).$$

Clearly, $\mathbb{E}_x[\tau(D)|\tau(D) < \tau(y)] \leq \mathbb{E}_x[\tau(y)|\tau(D) < \tau(y)]$, so

$$\begin{aligned}
\mathbb{E}_x[\tau(D)] &\leq \mathbb{E}_x[\tau(y)|\tau(D) < \tau(y)]q_x(D) + \mathbb{E}_x[\tau(y) \mid \tau(D) > \tau(y)]q_x(y) + \mathbb{E}_y[\tau(D)]q_x(y) \\
&= \mathbb{E}_x[\tau(y)] + \mathbb{E}_x[\tau(D)]q_x(y) \leq \mathbb{E}_x[\tau(y)] + \mathbb{E}_y[\tau(D)].
\end{aligned}$$

Hence,

$$\mathbb{E}_x[\tau(D)] - \mathbb{E}_y[\tau(D)] \leq \mathbb{E}_x[\tau(y)].$$

By switching x and y the same calculation also yields

$$\mathbb{E}_y[\tau(D)] - \mathbb{E}_x[\tau(D)] \leq \mathbb{E}_y[\tau(x)].$$

Now we consider two cases.

1. Assume $\mathbb{E}_x[\tau(D)] > \mathbb{E}_y[\tau(D)]$. Then,

$$\frac{|\mathbb{E}_x[\tau(D)] - \mathbb{E}_y[\tau(D)]|}{\max\{\mathbb{E}_x[\tau(D)], \mathbb{E}_y[\tau(D)]\}} = \frac{\mathbb{E}_x[\tau(D)] - \mathbb{E}_y[\tau(D)]}{\mathbb{E}_x[\tau(D)]} \leq \frac{\mathbb{E}_x[\tau(y)]}{\mathbb{E}_x[\tau(D)]}.$$

2. Is $\mathbb{E}_x[\tau(D)] > \mathbb{E}_y[\tau(D)]$, we have

$$\frac{|\mathbb{E}_x[\tau(D)] - \mathbb{E}_y[\tau(D)]|}{\max\{\mathbb{E}_x[\tau(D)], \mathbb{E}_y[\tau(D)]\}} = \frac{\mathbb{E}_y[\tau(D)] - \mathbb{E}_x[\tau(D)]}{\mathbb{E}_y[\tau(D)]} \leq \frac{\mathbb{E}_y[\tau(x)]}{\mathbb{E}_y[\tau(D)]}.$$

Putting both cases together proves the assertion. □

Note that Theorem 3 (i) gives the triangle inequality for the distance $d(x, y) = \mathbb{E}_x[\tau(y)]$ if we also consider D to consist of a single element. So $d(x, y)$ describes an asymmetric distance measure and the symmetrization $m(x, y) = 1/2(d(x, y) + d(y, x))$ is a metric.
Theorem 3 (ii) has an interesting consequence in the case that $x, y \in M$ are elements of a metastable set M and D is an arbitrary test set with $D \cap M = \emptyset$. In this case, both ratios

$$\frac{\mathbb{E}_x[\tau(y)]}{\mathbb{E}_x[\tau(D)]} \quad \text{and} \quad \frac{\mathbb{E}_y[\tau(x)]}{\mathbb{E}_y[\tau(D)]}$$

should be very small, since the expected time to travel between elements within a metastable set M should be much smaller than the expected time to leave the metastable set and travel to another set D. Then, Theorem 3 implies that the hitting time function

$$m_D(x) = \mathbb{E}_x[\tau(D)]$$

should be almost constant on metastable sets, and this property should be robust against the choice of the test set D.

4.3 Lower bounds using core sets

Next we want to add a quantitative estimate to our more qualitative argument regarding the constancy of the exit times on metastable sets. To this end let us study the exit time function

$$f(x) = \mathbb{E}_x[\tau(A^c)]$$

from a metastable set A. There may be states in A that belong to the transition region, i.e., for which $f(x)$ is not really large, perhaps the function f is not even constant in the neighborhood of these states. However, assume that there is a core set $C \subset A$ with the property that C contains most of the measure of A, and the communication of states in C is much quicker than the exit from A, i.e.,

(A1) there is a small $\epsilon > 0$ s.t. $\frac{\mu(C)}{\mu(A)} = 1 - \epsilon$, and

(A2) there are constants $\delta > 0$ and $M > 1$ with $\delta/M \ll 1$ s.t.

  (a) $\mathbb{E}_x(\tau(y)) \leq \delta$ for $x, y \in C$, and
  (b) $f(x) \geq M$ for all $x \in C$.

Then, we find the following statement regarding the value of the quantity

$$\hat{\kappa}(A) = \frac{1}{f_{max}} \frac{1}{\mu(A)} \sum_{x\in A} f(x)\mu(x), \qquad f_{max} = \max_{x\in A} f(x).$$

associated with A in Theorem 2:

Theorem 4 *Assume that the set A satisfies the conditions (A1) and (A2) given above for a specific set $C \subset A$. Furthermore assume that there is a constant $K > 0$ such that for all $x \in A \setminus C$ we have*

$$\max_{x\in A\setminus C} \left\{ \frac{\mathbb{E}_x[\tau(y)]}{f(x)}, \frac{\mathbb{E}_y[\tau(x)]}{f(y)} \right\} \le K,$$

for $y \in A$ with $f(y) = f_{max}$, and $K = 0$ if $A = C$. Then

$$1 - \hat{\kappa}(A) \le \frac{\delta}{M}(1-\epsilon) + K\epsilon,$$

such that $1 - \hat{\kappa}(A) \ll 1$ as long as $K\epsilon \ll 1$. Moreover we get the estimate

$$1 - \hat{\lambda}(A) \le \frac{1}{M(1-\epsilon)}.$$

for

$$\hat{\lambda}(A) = 1 - \frac{1}{e(A)}, \quad e(A) = \mathbb{E}_A(\tau(A^c)).$$

Proof. From Thm. 3 we get with $D = A^c$ for all $x \in A$ that

$$\frac{f_{max} - f(x)}{f_{max}} \le \max \left\{ \frac{\mathbb{E}_x[\tau(y)]}{f(x)}, \frac{\mathbb{E}_y[\tau(x)]}{f_{max}} \right\}.$$

which implies that

$$\begin{aligned} 1 - \hat{\kappa}(A) &= \frac{1}{f_{max}} \left[\frac{1}{\mu(A)} \sum_{x\in A} (f_{max} - f(x))\, \mu(x) \right] \\ &\le \frac{1}{\mu(A)} \sum_{x\in A} \max \left\{ \frac{\mathbb{E}_x[\tau(y)]}{f(x)}, \frac{\mathbb{E}_y[\tau(x)]}{f_{max}} \right\} \mu(x). \end{aligned}$$

For $x \in C$ condition (A2) yields

$$\max \left\{ \frac{\mathbb{E}_x[\tau(y)]}{f(x)}, \frac{\mathbb{E}_y[\tau(x)]}{f_{max}} \right\} \le \frac{\delta}{M},$$

while for $x \in A \setminus C$ we have

$$\max_{x\in A\setminus C} \left\{ \frac{\mathbb{E}_x[\tau(y)]}{f(x)}, \frac{\mathbb{E}_y[\tau(x)]}{f_{max}} \right\} \le K.$$

With these estimates we get with condition (A1) that

$$1 - \hat{\kappa}(A) \le \frac{1}{\mu(A)} \left(\mu(C)\frac{\delta}{M} + K\mu(A \setminus C) \right) = \frac{\delta}{M}(1-\epsilon) + K\epsilon.$$

Using condition (A2) again we find for $e(A) = (\sum_{x\in A} f(x)\mu(x))/\mu(A)$ that

$$e(A) \ge \frac{1}{\mu(A)}(M\mu(C) + \mu(A \setminus C) \cdot 1) = M(1-\epsilon) + \epsilon \ge M(1-\epsilon)$$

which implies the second assertion since $1 - \hat{\lambda}(A) = 1/e(A)$. □

This result shows that for a decomposition $\mathcal{D} = \{A_1, \ldots, A_m\}$ in which every set A_i contains a core set C_i such that $\mu(C_i) \approx \mu(A_i)$, and mixing within C is fast compared to exit times from A_i if starting from C_i, then $\mathcal{D}$ is a metastable partition in the sense that the metastability index $\mathcal{M}(\mathcal{D})$ is very close to 1.

Example: Let us consider the 7-state Markov chain with the following transition matrix:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0.99 & 0 & 0 & 0.01 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.02 & 0 & 0.98 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

The chain is non-reversible with invariant measure

$$\mu^T = (0.2210, 0.2210, 0.2232, 0.0045, 0.1116, 0.1094, 0.1094).$$

There clearly are two metastable sets with state 4 being a kind of transition state that belongs more to state 5-7 than to 1-3. We therefore choose $A_1 = \{1, 2, 3\}$ and $A_2 = \{4, 5, 6, 7\}$. This choice leads to the metastability index

$$\mathcal{M}(\{A1, A2\}) = 0.995.$$

Furthermore direct computation leads to $\hat{\lambda}_1 = 0.997$, $\hat{\kappa}_1 = 0.997$, $\hat{\lambda}_2 = 0.997$, and $\hat{\kappa}_2 = 0.990$, so that the lower bound to the metastability index as of Thm. 2 results as

$$\frac{1}{2}\left[\hat{\lambda}_1\hat{\kappa}_1 + \hat{\lambda}_2\hat{\kappa}_2\right] = 0.990.$$

In order to compute the lower bound resulting from Thm. 4 we choose the core sets $C_1 = \{1, 2, 3\}$ and $C_2 = \{5, 6, 7\}$ such that $\delta(A_1) = 3.53$, $M(A_1) = 298$, $\delta(A_2) = 8.14$, $M(A_2) = 298$, $\epsilon(A_2) = 0.013$, and $K(A_2) = 2.05$ which results in the estimate

$$\frac{1}{2}\left[\hat{\lambda}_1\hat{\kappa}_1 + \hat{\lambda}_2\hat{\kappa}_2\right] \geq 0.982.$$

4.4 Algorithmic considerations

The spectral result stated in Theorem 1 can be exploited to algorithmically identify a metastable partition of a reversible Markov chain. First, we need to calculate the dominant spectrum, so the eigenvalues $1 = \lambda_1 > \lambda_2 \geq ... \geq \lambda_n$ that are close to 1 and the corresponding eigenvectors $u_1, ..., u_n$. Then, we want to find a partition of state space into sets such that the dominant eigenvectors are as constant as possible on these sets. As mentioned above, this problem can be ill-conditioned due to the existence of states in a transition region that cannot be clearly assigned to a certain metastable set. The transition region should be a set of states that the process typically leaves quickly into one of the core sets [19]. In [17] it was discussed how -for reversible chains- we can first identify the transition region such that we can next restrict the problem to partitioning the core sets only: we assume that we are provided with a reference measure μ_r that allows us to find the tranation region via

$$\mathcal{T} = \{x \in \mathbb{X} \mid (P^T)^\alpha \mu_r(x) < \mu_r(x)\},$$

where $\alpha \in \mathbb{N}$ is the approximate average timescale on which the transition region is left by the Markov chain. Such a measure exists, e.g., when the Markov chain depends on a parameter that controls the metastability of the process: we have the original Markov chain with transition matrix P and another Markov chain with reduced metastability, transition matrix P_r, and unique stationary distributions μ and μ_r, see, e.g., [17] for random walks on networks, or [19] for molecular dynamics with temperature embedding.

If the Markov chain is non-reversible, we will claim that the transition region need to be left quickly either forward or backward in time. That is, assuming $\mu > 0$ we can define the transition matrix P_b of the time-reversed Markov chain with entries

$$p_b(x, y) = \frac{\mu(y)}{\mu(x)} p(y, x)$$

and the transition region as

$$\mathcal{T} = \{x \in \mathbb{X} \mid (P^T)^\alpha \mu_r(x) < \mu_r(x) \text{ or } (P_b^T)^\alpha \mu_r(x) < \mu_r(x)\}. \tag{2}$$

This is equivalent to claiming that the core sets are attractive sets in forward and backward time. If we do not have a parameter for controlling the metastability of the process, we simply choose μ_r to be a uniform distribution in our finite state space.

Having restricted the clustering problem to $C = \mathbb{X} \setminus \mathcal{T}$, the state space without transition region, there exist optimization-based algorithmic strategies for finding partitions on which the dominant eigenvectors are almost constant [9, 12, 10]. Another idea is to associate with every state $x \in C$ the vector $u(x) = (u_1(x), ..., u_n(x)) \in \mathbb{R}^n$ and apply aa appropriate cluster analysis technique to the points $(u(x))_{x \in C}$ in $\mathbb{R}^n$, for example the k-means algorithm. In the general case of a non-reversible Markov chain, the eigenvalues and eigenvectors might not be real-valued and it is not clear how to use this spectral information for clustering in this case. In the next section, we will further discuss this issue using a simple network example. Our mathematical results motivate to replace the m leading eigenvectors by m hitting time functions with respect to suitable test sets. "Suitable" here means that the test sets should not lead to redundant information in terms of hitting times, for example, the sets should not be too close to each other. So the idea is to compute for n sets $D_1, ..., D_n$ the hitting time functions $m_i = m_{D_i}$ and use the same algorithmic strategies [9, 12, 10] on the functions m_i instead of the eigenvectors u_i. We will later also use a k-means clustering of the points $(m(x))_{x \in C}$ in $\mathbb{R}^n$ with $m(x) = (m_1(x), ..., m_n(x))$. Note that for such approaches the hitting time functions should be normalized as the eigenvectors are. If the transition matrix P of the Markov chain is known, a hitting time function m_i can be computed as the solution of the following linear system

$$\begin{aligned} (P - Id)m_i &= -1 \qquad \text{on } \mathbb{X} \setminus D_i, \\ m_i &= 0 \qquad\qquad \text{on } D_i. \end{aligned}$$

Moreover, the value of a hitting time function $m_i(x) = \mathbb{E}_x[\tau(D_i)]$ at a certain point $x \in \mathbb{X}$ could also be estimated by Monte Carlo simulation. In this sense a hitting time function can also be evaluated locally. These two properties clearly yield computational advantages over the calculation of the dominant eigenvectors, even in the reversible case.

Summarizing, we find the following algorithmic blueprint:

1. We choose a number n of test sets and a metastability parameter $\alpha \in \mathbb{N}$.
2. Given P, P_b, and μ_r we compute the transition region $\mathcal{T}$ as in (2).
3. We consider a test set $D_1 = \{x_1\}$ consisting of a single state $x_1 \in C = \mathbb{X} \backslash \mathcal{T}$ that is chosen randomly and compute the hitting time function m_1.
4. For $i = 2, ..., n$ we choose the test set $x_i = \arg\max_{x \in C} \min_{j=1,..,i-1} m_j(x)$ and compute the hitting time function m_i with respect to $D_i = \{x_i\}$.
5. We normalize the functions m_i and fully partition C by clustering the points $m(x) = (m_1(x), ..., m_n(x)) \in \mathbb{R}^n$ with $x \in C$, for example using k-means.

5 Numerical Examples

In this section, we will use our algorithmic strategy to identify metastable sets for a random walk on a network, see Section 2.2. First, let us illustrate the relation between the spectral and hitting time based approach using an undirected network that is shown in Figure 2 where the induced random walk is a reversible Markov chain.

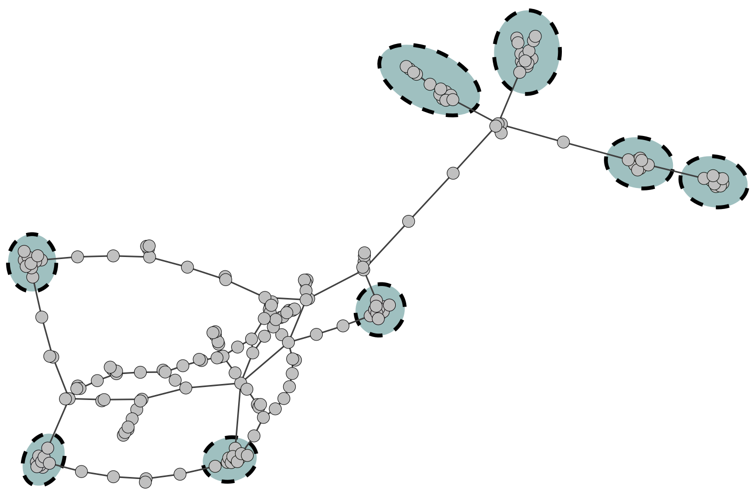

Figure 2: Undirected network with 8 strongly connected modules and a loosely connected transition region.

5.1 Reversible chain

The first step of the algorithm identifies the transition region as in Fig. 2 and in Fig. 3 the values of the three dominant eigenvectors of the random walk on the core sets $C = \mathbb{X} \setminus \mathcal{T}$ are shown in comparison to three hitting time functions with respect to test sets generated by the algorithm. Here we restrict the functions to the region $C = \mathbb{X} \setminus \mathcal{T}$ that we want to partition.

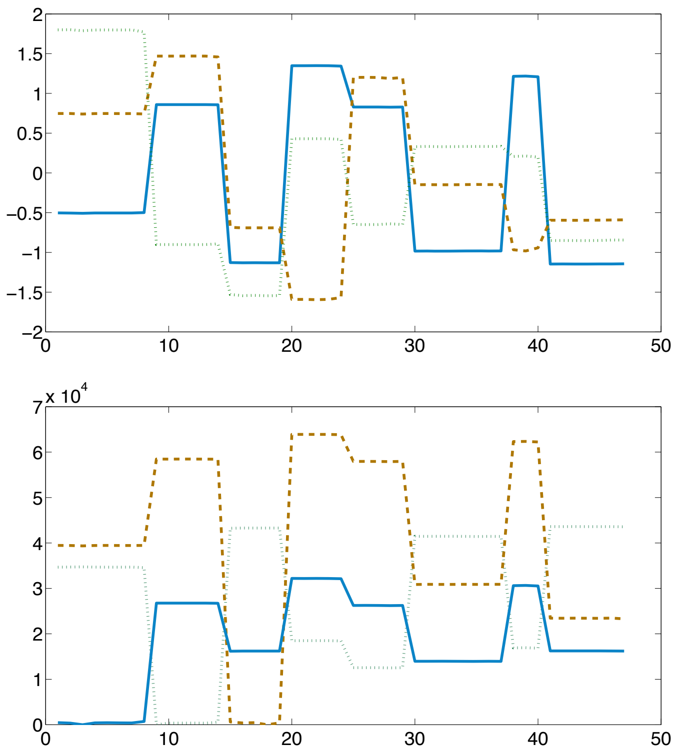

Figure 3: Upper panel: The 3 dominant non-trivial eigenvectors, shown in the reduced state space $C = \mathbb{X} \setminus \mathcal{T}$. Lower panel: 3 hitting time functions in C as generated by the algorithm.

As expected from Thm. 1 and Thm. 3 the eigenvectors and hitting time functions are almost constant on the metastable sets. The k-means method described in the previous section delivers in both cases the clustering that is visualized in Figure 4.

5.2 Non-reversible chain

Our second example should underline the difficulty of extending spectral methods to non-reversible cases where loops and cyclic structures are present. We consider the directed network that is illustrated in Figure 5.

The network consists of 3 cycles of length 5 that are connected via one transition node. Additionally, the two upper cycles in Fig. 5 share one undirected edge. When applying the proposed method based on hitting time functions for

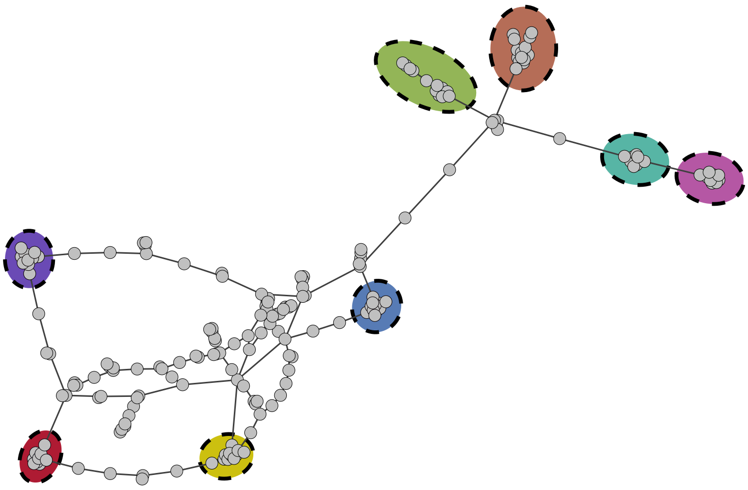

Figure 4: Identification of 8 modules by k-means using an embedding of 3 eigenvectors or 3 hitting time functions into $\mathbb{R}^3$.

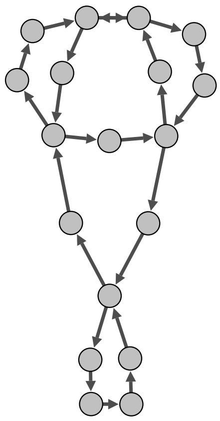

Figure 5: Directed network with 3 cycles that are metastable for the standard random walk.

finding metastable cluster in the network, we get the result for 2 and 3 sets as illustrated in Fig. 6.

If we calculate the values

$$\hat{\lambda}_i = 1 - \frac{1}{\mathbb{E}_{A_i}[\tau(A_i^c)]}$$

from Theorem 2 for the three cycles that had been identified as cluster as shown in Fig. 6, we find

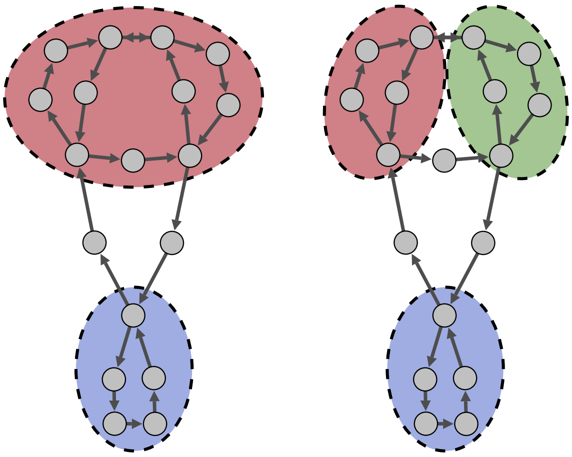

Figure 6: Left: Clustering into 2 sets. Right: Clustering into 3 sets.

i	$\hat{\lambda}_i$	color
1	0.7529	red
2	0.7529	green
3	0.8696	blue

On the other hand, the eigenvalues of the random walk on the network are complex and the eigenvalues with the largest real part are given by

i	λ_i
1	1
2	0.7156 + 0.0053i
3	0.7156 - 0.0053i
4	0.3593 + 0.8579i
5	0.3593 - 0.8579i
6	0.3260 + 0.6374i
7	0.3260 - 0.6374i

That is, we have one complex conjugate pair of eigenvalues that has a dominating real part that is at least not far away from 1; this agrees with the metastability of the 3 cycles. The following table shows the values of the real part of the dominant eigenvector on the cluster and the transition region found by the hitting time apporach.

Real part of first non-trivial eigenvector on blue cluster

x	1	2	3	4	5
$Re(u_2(x))$	-0.0534	-0.0196	0.0035	0.0185	-0.1011

Real part of first non-trivial eigenvector on red cluster

x	6	7	8	9	10
$Re(u_2(x))$	0.3340	0.2626	0.2039	0.1561	0.1177

Real part of first non-trivial eigenvector on green cluster

x	11	12	13	14	15
$Re(u_2(x))$	-0.1596	-0.1362	-0.2329	-0.2092	-0.1843

Real part of first non-trivial eigenvector on transition region

x	16	17	18
$Re(u_2(x))$	-0.2357	-0.2357	-0.2357

We can see that the sign structure [9, 12] of the real part of the eigenvector does not coincide with the clustering. In particular, the eigenvector changes its sign on the most metastable set (blue in Fig. 6) and achieves its minimal value $\min Re(u_2(x)) = -0.2357$ in the transition region, where for reversible Markov chains typically values close to 0 are expected.

This example emphasizes the difficulty of a metastable spectral clustering for non-reversible Markov chains that show a strongly cyclic behavior within and between metastable sets. On the other hand, as motivated by the theoretical results and demonstrated by this example our developed approach based on hitting time analysis is well applicable in such cases.

6 Conclusion

In this article, we discussed the relation of metastable or almost invariant sets of a Markov chain to spectral properties of its transition matrix on the one hand and to properties of hitting and escape times on the other. While the spectral approach is restricted to reversible Markov chains, in general, the hitting time approach proposed herein can also be applied to the non-reversible case.

When wanting to find modules or cluster in a network this extension is necessary since for a network the associated random walk is only reversible if the network is undirected. Even if the Markov chain is reversible and both approaches are applicable we highlighted computational advantages of the hitting time approach.

Finally, we illustrated the theoretical results and algorithmic ideas on two network examples. The first example was undirected, underlining the close relation of the two approaches in the reversible case. The second example was a directed network with loops and strongly cyclic structures where the spectral information does not allow for identifiyng the metastable partition that can well be computed from appropriate hitting time functions in the way demonstrated.

References

[1] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and small eigenvalues in Markov chains. *J. Phys. A*, 33(46):L447–L451, 2000.

[2] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and low lying spectra in reversible Markov chains. *Comm. Math. Phys.*, 228:219–255, 2002.

[3] 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.

[4] G. E. Cho and C. D. Meyer. Aggregation/disaggregation methods for nearly uncoupled Markov chains. *Technical Report NCSU no. 041600-0400, North Carolina State University*, 1999.

[5] M. Dellnitz, G. Froyland, and O. Junge. The algorithms behind GAIO - set oriented numerical methods for dynamical systems. In *Ergodic theory, analysis, and efficient simulation of dynamical systems*, pages 145–174. Springer, 2001.

[6] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. *SIAM J. Num. Anal.*, 36(2):491–515, 1999.

[7] M. Dellnitz and O. Junge. Set oriented numerical methods for dynamical systems. In *Handbook of dynamical systems*, pages 221–264. North-Holland, Amsterdam, 2002.

[8] M. Dellnitz and R. Preis. Congestion and almost invariant sets in dynamical systems. In *Symbolic and Numerical Scientific Computation*, pages 183–209. LNCS, 2003.

[9] P. Deuflhard, W. Huisinga, A. Fischer, and C. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Lin. Alg. Appl.*, 315:39–59, 2000.

[10] P. Deuflhard and M. Weber. Robust Perron cluster analysis in conformation dynamics. *Linear Algebra and its Applications*, 161(184), 2005. 398 Special issue on matrices and mathematical biology.

[11] D. Fritzsche, V. Mehrmann, D. B. Szyld, and E. Virnik. An SVD approach to identifying meta-stable states of Markov chains. *Electronic Transactions on Numerical Analysis*, 29:46–69, 2008.

[12] W. Huisinga and B. Schmidt. Metastability and dominant eigenvalues of transfer operators. In C. Chipot, R. Elber, A. Laaksonen, B. Leimkuhler, A. Mark, T. Schlick, C. Schütte, and R. Skeel, editors, *Advances in Algorithms for Macromolecular Simulation*, Lecture Notes in Computational Science and Engineering 49. Springer, 2002.

[13] S. Lafon and A. Lee. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:1393–1403, 2006.

[14] T. Li, W. E, and E. Vanden Eijnden. Optimal partition and effective dynamics of complex networks. *Proc. Nat. Acad. Sci.*, 105, 2008.

[15] E. Meerbach, C. Schütte, and A. Fischer. Eigenvalue bounds on restrictions of reversible nearly uncoupled Markov chains. *Lin. Alg. Appl.*, 398, 2005.

[16] C. D. Meyer. Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems. *SIAM Rev.*, 31, 1989.

[17] M. Sarich, N. Djurdjevac, S. Bruckner, T. O. F. Conrad, and C. Schütte. Modularity revisited: A novel dynamics-based concept for decomposing complex networks. *Journal of Computational Dynamics*, 2013.

[18] C. Schütte, A. Fischer, W. Huisinga, and P. Deuflhard. A direct approach to conformational dynamics based on hybrid Monte Carlo. *J. Comput. Phys.*, 151:146–168, 1999. Special Issue on Computational Biophysics.

[19] C. 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.