3 Metastability

We present a mathematical characterization of metastability, its connection to eigenvalues of the transfer operator close to its dominant eigenvalue 1, and a theoretical as well as numerical identification strategy.

3.1 Characterizing Metastability

There is no unique but several definitions of metastability in literature (see, e.g., [4, 9, 72, 73]), each adapted to suit the context. Our definition fits the biomolecular application context and measures metastability w.r.t. the canonical ensemble. Combining the physical characterization of metastable sub–ensembles in Section 1.1 and the mathematical specification in Section 1.3, we aim at defining a transition probability from a subset B into C within the time span τ, denoted by $p(\tau, B, C)$, such that an invariant sub–ensemble C is characterized by $p(\tau, C, C) = 1$, while a metastable sub–ensemble can be characterized by $p(\tau, C, C) \approx 1$. Being "close to 1" is obviously a vague statement—however, in most applications we are interested in a decomposition into the most metastable subsets, which eliminates the problem of interpreting "close to 1". Instead we have to determine the number of subsets, we are looking for. In our approach, this is done by examining the spectrum of the propagator P_τ. Alternatively, we could determine a cascade of decompositions with an increasing number of metastable subsets.

Define the **transition probability** $p(t, B, C)$ from $B \in \mathcal{A}$ to $C \in \mathcal{A}$ within the time span t as the conditional probability

$$p(t, B, C) \;=\; \mathbf{P}_\mu[X_t \in C \;|X_0 \in B] \;=\; \frac{\mathbf{P}_\mu[X_t \in C \text{ and } X_0 \in B]}{\mathbf{P}_\mu[X_0 \in B]}, \qquad (31)$$

where $\mathbf{P}_\mu$ indicates that the initial distribution of the Markov process X_t is due to μ, hence $X_0 \sim \mu$. The similar symbols for both the transition probability between subsets $p(t, B, C)$ as well as for the stochastic transition function corresponding to the Markov process emphasize the strong relation to the definition of $p(t, x, C)$ in (4), which allows to rewrite (31) as

$$p(t, B, C) \;\;=\;\; \frac{1}{\mu(B)} \int_B p(t, x, C)\, \mu(\mathrm{d}x). \qquad (32)$$

The transition probability quantifies the dynamical fluctuations within the stationary ensemble μ. For later reference, we state the following two properties:

(i) using the duality bracket $\langle \cdot, \cdot \rangle_\mu$ between $L^1(\mu)$ and $L^\infty(\mu)$, we get

$$p(t, B, C) \;\;=\;\; \frac{\langle P_t \mathbf{1}_B, \mathbf{1}_C \rangle_\mu}{\langle \mathbf{1}_B, \mathbf{1}_B \rangle_\mu}. \qquad (33)$$

(ii) metastability of C may equivalently be characterized by the condition that $p(t, C, \mathbf{X} \setminus C) \approx 0$, which yields [73]:

$$p(t, C, \mathbf{X} \setminus C) \;=\; \frac{1}{2\mu(C)} \left\| P_t \mathbf{1}_C \,-\, \mathbf{1}_C \right\|_1 . \tag{34}$$

In Section 1.1 we have seen that metastability of sub–ensembles can experimentally be measured w.r.t. some observation time τ. Therefore, we will fix some $\tau > 0$ and concentrate in the sequel on the single propagator P_τ rather than on the entire semigroup of propagators $\{P_t\}_{t \in \mathbf{T}}$.

3.2 Identifying Metastable Subsets

We now present the fundamental algorithmic strategy used to identify metastable subsets. The basic idea is to interpret metastability as almost invariance. Since invariant subsets are associated with the eigenvalue $\lambda = 1$ and can be identified exploiting the corresponding eigenfunctions [17], metastable subsets are thought to be associated with almost $\lambda = 1$ eigenvalues and can be identified by exploiting the corresponding eigenfunctions.

Consider the propagator $P_\tau : L^r(\mu) \to L^r(\mu)$ with $r = 1, 2$; its spectrum is contained in the unit disc $\{\lambda \in \mathbf{C} : |\lambda| \leq 1\}$. Whenever a proper subset $C \subset \mathbf{X}$ is invariant under the Markov process, i.e., $p(t, x, \mathbf{X} \setminus C) = 0$ for all $x \in C$, the probability density $\hat{\mathbf{1}}_C = \mathbf{1}_C / \mu(C)$ is an eigenfunction corresponding to $\lambda = 1$. In particular, since μ is assumed to be invariant, $\hat{\mathbf{1}}_{\mathbf{X}}$ is an eigenfunction corresponding to $\lambda = 1$. Loosely speaking, a characterization of metastability according to (34) suggests that C is metastable if $\hat{\mathbf{1}}_C$ is an approximate eigenfunction corresponding to an eigenvalue close to $\lambda = 1$. This motivates the following **algorithmic strategy**:

> Metastable subsets (on the time scale $\tau > 0$) can be identified via eigenfunctions of the propagator P_τ corresponding to eigenvalues $|\lambda| < 1$ close to the Perron root $\lambda = 1$. In doing so, the number of metastable subsets is equal to the number of eigenvalues close to 1, including $\lambda = 1$ and counting multiplicity.

The strategy mentioned above has first been proposed by Dellnitz and Junge [13] for discrete dynamical systems with weak random perturbations and has been successfully applied to molecular dynamics in different contexts [69, 71, 68]; a justification is given by Theorem 3.1. The algorithmic strategy necessitates the following two **conditions on the propagator** P_τ:

(C1) The essential spectral radius of P_τ is less than one, i.e., $r_{\text{ess}}(P_\tau) < 1$.

(C2) The eigenvalue $\lambda = 1$ of P_τ is simple and dominant, i.e., $\eta \in \sigma(P_\tau)$ with $|\eta| = 1$ implies $\eta = 1$.

It is important to remember that we defined transition probabilities between subsets, and therefore metastability, w.r.t. the invariant measure μ. Assume that the Markov process X_t admits another invariant measure ν, which, for sake of simplicity, is absolutely continuous w.r.t. μ with density $f \in L^1(\mu)$. Then, f is an eigenfunction of P_τ corresponding to $\lambda = 1$. As a consequence, we will not be able to decide in general whether an eigenfunction corresponding to some eigenvalue $\lambda \approx 1$ is related to metastable behavior of the ensemble represented by μ or by ν. Thus, the algorithmic strategy requires *uniqueness* of the invariant measure. Additionally, the physical interpretation of the ensemble excludes other eigenvalues than $\lambda = 1$ *on* the unit circle. Hence, $\lambda = 1$ has to be simple and dominant. For the numerical realization and discretization of the eigenvalue problem, we moreover need that the relevant eigenvalues are *isolated and of finite multiplicity*. For those eigenvalues convergence results of the numerical discretization algorithm can be established. This implies that the essential spectral radius has to be less than 1, hence permitting the existence of isolated eigenvalues of finite multiplicity close to $\lambda = 1$.

3.3 Metastable Subsets and Eigenvalues Close to 1

We now give a mathematical justification of the algorithmic strategy introduced above. The main result is stated in Theorem 3.1. It illuminates the strong relation between the existence of a cluster of eigenvalues close to 1 and a possible decomposition of the state space into metastable subsets. We state the theorem under the additional assumption of reversibility of the Markov process X_t and comment on how the results can be applied to non–reversible Markov processes.

Consider the propagator $P_\tau : L^2(\mu) \to L^2(\mu)$ satisfying the two conditions (C1) and (C2), and assume that the Markov process is reversible. Due to Proposition 1.1, P_t is self–adjoint and its spectrum has the form

$$\sigma(P_\tau) \subset [l, r] \cup \{\lambda_2\} \cup \{1\},$$

with $-1 < l \leq r \leq \lambda_2 < \lambda_1 = 1$. We restrict our considerations to the case that the Perron root is "nearly two–fold degenerate": We assume that λ_2 is a simple isolated eigenvalue, hence $r = \lambda_3 < \lambda_2$, and further that the corresponding eigenfunction v_2 is normalized by $\langle v_2, v_2 \rangle_\mu = 1$ and satisfies $v_2 \in L^\infty(\mu)$. Note that $\langle v_2, \mathbf{1}_\mathbf{X} \rangle_\mu = 0$, since P_τ is self–adjoint.

A **decomposition** $\mathcal{D} = \{D_1, \ldots, D_n\}$ of the state space $\mathbf{X}$ is a collection of subsets $D_k \subset \mathbf{X}$ with the properties:

(i) positivity: $\mu(D_k) > 0$ for every k,

(ii) disjointness: $D_k \cap D_l = \emptyset$ for $k \neq l$, and

(iii) covering property: $\cup_{k=1}^{n} D_k = \mathbf{X}$.

For a decomposition $\mathcal{D} = \{B, C\}$ of $\mathbf{X}$ into two subsets, we define the following function

$$v_{BC} \;=\; \sqrt{\frac{\mu(C)}{\mu(B)}}\,\mathbf{1}_B \;-\; \sqrt{\frac{\mu(B)}{\mu(C)}}\,\mathbf{1}_C\,, \tag{35}$$

which is constant on either of the two sets B and C, and is normalized to $\|v_{BC}\|_2 = 1$. Under the assumptions on the propagator P_τ stated above we obtain the following relation between the existence of metastable subsets and eigenvalues close to 1.

Theorem 3.1 *Let $\mathcal{D} = \{B, C\}$ be an arbitrary decomposition of $\mathbf{X}$ into two subsets. Then*

$$1 + \kappa\lambda_2 \;\;\leq\;\; [\, p(\tau, B, B) + p(\tau, C, C)\,] \;\leq\; 1 + \lambda_2, \tag{36}$$

with $\kappa = \langle v_2, v_{BC}\rangle_\mu^2 \leq 1$. In addition, choosing

$$B = \{x \in \mathbf{X} : v_2(x) \geq 0\} \quad \textit{and} \quad C = \{x \in \mathbf{X} : v_2(x) < 0\},$$

we have $1 - 8c^2\epsilon \leq \kappa$ with constants $\epsilon = (1 - \lambda_2)/(1 - \lambda_3)$ and $c = \|v_2\|_\infty$.

Proof: The proof is based on results by Davies [9, 10] and a subsequent paper of Singleton [73]. In order to be applicable to our situation, we have to extend their results, since in general, we cannot assume the existence of an infinitesimal generator as in [9, 10, 73]. In [73] we have to replace the strongly continuous semigroup $\exp(-Ht)$ by P_τ. Furthermore, to match the assumptions on the spectrum in [73] with ours, we have to rescale the time of the semigroup. Interpreting $p(\tau, B, C)$ as $\mu(C)\langle P_\tau \mathbf{1}_B/\mu(B), \mathbf{1}_C/\mu(C)\rangle_\mu$, as stated by property (i) in Section 3.1, Lemma 4 of [73] and its subsequent remark state that

$$p(\tau, B, C) \;\;=\;\; \frac{1}{2\mu(B)}\|\mathbf{1}_B - P_\tau \mathbf{1}_B\|_1 \;=\; \mu(C)\,\langle v_{BC} - P_\tau v_{BC}, v_{BC}\rangle_\mu.$$

We then exploit Theorem 5 of [73] to bound the scalar product by

$$1 - \lambda_2 \;\;\leq\;\; \langle v_{BC} - P_\tau v_{BC}, v_{BC}\rangle_\mu \;\leq\; 1 - \kappa\lambda_2$$

with $\kappa = \langle v_2, v_{BC}\rangle_\mu^2 \leq 1$. Putting everything together, we end up with

$$\mu(B)(1 - \lambda_2) \;\;\leq\;\; p(\tau, B, C) \;\leq\; \mu(B)(1 - \kappa\lambda_2). \tag{37}$$

Repeating the calculation with $v_{CB} = -v_{BC}$ and exchanged roles of B and C, we see that inequality (37) holds in an analogous way for $p(\tau, C, B)$.

Hence, summing up both inequalities and exploiting $\mu(B) + \mu(C) = 1$, this finally gives

$$\begin{aligned} 1 - \lambda_2 \;&\leq\; [\, p(\tau, B, C) + p(\tau, C, B)\,] \;\leq\; 1 - \kappa\lambda_2 \\ \Leftrightarrow \quad 1 + \kappa\lambda_2 \;&\leq\; [\, p(\tau, B, B) + p(\tau, C, C)\,] \;\leq\; 1 + \lambda_2, \end{aligned}$$

which is the first statement of Theorem 3.1. For the second statement on the lower bound on κ, we conclude from Theorem 3 of [73] that

$$\kappa \;=\; 1 - \frac{1}{2}\|v_2 - v_{BC}\|_2^2. \tag{38}$$

In the following we determine, analogous to Theorem 3 in [9], an upper bound on $\|v_2 - v_{BC}\|_2$, which implies a lower bound on κ. Once again, $\exp(-Ht)$ has to be replaced by P_τ and furthermore, the infinitesimal generator $-H$ has to be substituted by $\mathrm{Id} - P_\tau$. Define the function

$$\psi \;=\; \sqrt{\frac{\|v_2^-\|_2}{\|v_2^+\|_2}}\, v_2^+ + \sqrt{\frac{\|v_2^+\|_2}{\|v_2^-\|_2}}\, v_2^-,$$

where v_2^+ and v_2^- denote the positive and negative part of v_2, respectively. Note that $\langle P_\tau\psi, \psi\rangle_\mu \geq \lambda_1$ implies

$$\langle (\mathrm{Id} - P_\tau)\psi, \psi\rangle_\mu \;\geq 1 - \lambda_2. \tag{39}$$

Now, define $\xi = \psi - \langle \psi, \mathbf{1_X}\rangle \mathbf{1_X}$. Since ξ is orthogonal to $\mathbf{1_X}$ and v_2, we obtain by means of Cauchy–Schwarz $\langle P_\tau\xi, \xi\rangle \leq \lambda_3\|\xi\|_2^2$ and therefore

$$(1 - \lambda_3)\|\xi\|_2^2 \;\leq\; 1 - \lambda_3\|\xi\|_2^2 \;\leq\; 1 - \langle P_\tau\xi, \xi\rangle \;\leq\; \langle (\mathrm{Id} - P_\tau)\xi, \xi\rangle. \tag{40}$$

Assembling (39) and (40) results in

$$\begin{aligned} (1 - \lambda_3)\|\xi\|_2^2 \;&\leq\; \langle (\mathrm{Id} - P_\tau)\xi, \xi\rangle \\ &=\; \langle (\mathrm{Id} - P_\tau)\psi, \psi\rangle - \langle \psi, 1\rangle \langle (\mathrm{Id} - P_\tau)\psi, \mathbf{1_X}\rangle \\ &\quad\; - \langle \psi, 1\rangle \langle (\mathrm{Id} - P_\tau)\mathbf{1_X}, \psi\rangle + \langle \psi, 1\rangle^2 \langle (\mathrm{Id} - P_\tau)\mathbf{1_X}, \mathbf{1_X}\rangle \\ &=\; \langle (\mathrm{Id} - P_\tau)\psi, \psi\rangle \;\leq\; 1 - \lambda_2, \end{aligned}$$

which implies $\|\xi\|_2^2 \leq \epsilon$ with $\epsilon = (1 - \lambda_2)/(1 - \lambda_3)$. With this modification in the proof of Theorem 3 in [9], we finally get

$$\|v_2 - v_{BC}\|_2^2 \leq 16\|v_2\|_\infty^2 \epsilon,$$

which together with (38) gives the lower bound on κ. □

Theorem 3.1 highlights the strong relation between a decomposition of the state space into two metastable subsets and a second eigenvalue close

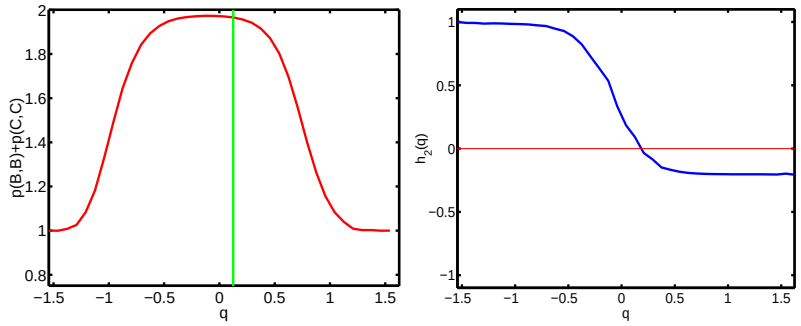

Figure 2: Data based on Smoluchowski equation. Left: metastability of a decomposition $\mathcal{D} = \{B, C\}$ with $B = (-\infty, q]$ and $C = (q, \infty]$ in dependence on q. The problem of finding the maximal value corresponding to the optimal decomposition is ill–conditioned. The vertical line corresponds to the decomposition $B = \{x \in \mathbf{X} : h_2(x) \geq 0\}$ and $C = \mathbf{X} \setminus B$ defined by the second eigenfunction h_2 (right). We obtain $\kappa = 0.984$ according to Theorem 3.1.

to the dominant eigenvalue 1. For an arbitrary decomposition $\mathcal{D} = \{B, C\}$ let us call $p(\tau, B, B) + p(\tau, C, C)$ the metastability of the decomposition $\mathcal{D}$. Then Theorem 3.1 states that the metastability of a decomposition $\mathcal{D} = \{B, C\}$ cannot be larger than $1 + \lambda_2$, while it is at least $1 + \kappa\lambda_2$. The upper bound is "large" whenever the eigenfunction v_2 corresponding to λ_2 is almost constant on the two metastable subsets B and C. As stated by the second part of Theorem 3.1, we can guarantee metastability for the particular decomposition into $B = \{x : v_2(x) \geq 0\}$ and $C = \{x : v_2(x) < 0\}$ whenever (i) the gap between the second and third eigenvalue is large, hence $\epsilon = (1 - \lambda_2)/(1 - \lambda_3)$ is small, and (ii) the essential maximum $c = \|v_2\|_\infty$ of the second eigenfunction v_2 is small. In [10] Davies proved, that in the case of a strongly continuous positive semigroup of self–adjoint propagators, e.g., in the case of the Smoluchowski dynamics, the lower bound on κ is in fact independent of $c = \|v_2\|_\infty$, whenever $c < \infty$. Nevertheless, the lower bound on $p(\tau, B, B) + p(\tau, C, C)$ via $\epsilon \ll 1$ implies the quite restrictive assumption: $\lambda_3 \ll \lambda_2$ on the spectrum P_τ. In numerical experiments we have observed intriguing results of the identification strategy even for situations corresponding to large ϵ-values [17].

In view of Theorem 3.1, it is natural to ask, whether there is an *optimal* decomposition with highest possible metastability. The answer is illustrated by Figure 2: *Even if there exists an optimal decomposition, the problem of finding it might be ill–conditioned.* The graph shows the metastability of a family of decompositions. It is based on the propagator P_τ corresponding to the Smoluchowski equation for the double–well potential on $\mathbf{X} = \mathbf{R}$. The conditions (C1) and (C2) on P_τ are justified by Proposition 6.5, while the assumption on the spectrum can be fulfilled by choosing an appropriate in-

verse temperature β. We identify a whole domain of decompositions that are nearly optimal. In this case the problem of finding the maximum is ill–conditioned. We also observe that the decomposition suggested by our identification algorithm is nearly optimal. The phenomenon illustrated by Figure 2 is believed to by typical in our application context, which is due to the fact that the canonical ensemble has large regions of almost vanishing probability.

Having an application to more complicated dynamical behavior in mind, we claim the following generalization of Theorem 3.1 for a decomposition into more than two subsets: Assume that the propagator P_τ, acting on $L^2(\mu)$, is associated with a reversible Markov process and satisfies the conditions (C1) and (C2). Moreover, assume that its spectrum is of the form $\sigma(P_\tau) \subset [l, r] \cup \{\lambda_n\} \cup \cdots \cup \{\lambda_2\} \cup \{1\}$ with simple, isolated eigenvalues $\lambda_n < \ldots < \lambda_2 < \lambda_1 = 1$ and corresponding eigenfunctions in $L^\infty(\mu)$. Given a decomposition $\mathcal{D} = \{D_1, \ldots, D_n\}$ of $\mathbf{X}$, denote by $\{v_{D_1}, \ldots, v_{D_n}\}$ some μ–orthonormal basis of span$\{\mathbf{1}_{D_1}, \ldots, \mathbf{1}_{D_n}\}$.

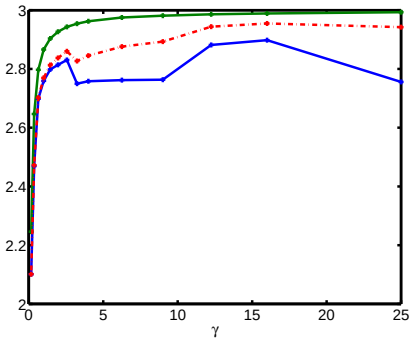

Figure 3: Illustration of the Conjecture for a decomposition of the state space into three subsets. The graph is based on the propagator P_τ corresponding to the Smoluchowski equation. The test system is the three–well potential defined in (63) for $\beta = 2$ and different values of γ. For details on the discretization see Section 6. The top solid line represents the upper bound $1 + \lambda_2 + \lambda_3$ in the Conjecture, while the dashed line corresponds to the metastability $p(\tau, D_1, D_1) + p(\tau, D_2, D_2) + p(\tau, D_3, D_3)$ of the decomposition obtained by applying the identification algorithm. The bottom solid line represents the lower bound $\kappa_1 1 + \kappa_2\lambda_2 + \kappa_3\lambda_3$.

Conjecture. Let $\mathcal{D} = \{D_1, \ldots, D_n\}$ be an arbitrary decomposition of $\mathbf{X}$ into n subsets. Then

$$\kappa_1 1 + \ldots + \kappa_n\lambda_n \leq [\, p(\tau, D_1, D_1) + \ldots + p(\tau, D_n, D_n)\,] \leq 1 + \ldots + \lambda_n$$

with $\kappa_j = \langle v_j, v_{D_j}\rangle_\mu^2 \leq 1$.

For a numerical verification of the Conjecture for a decomposition of the state space into $n = 3$ subsets see Figure 3. The conjecture is in agreement

with all numerical experiments performed in Section 6.

In the following, we want to comment on an extension to *non–reversible* Markov processes. In his PhD thesis [74], Singleton proved results about metastable states for non–self–adjoint strongly continuous Markov semigroups, which might be used to extend Theorem 3.1. The main problem in the absence of self–adjointness is to control both, the propagator and its adjoint at the same time. This necessitates increasing technical effort, since we have to pose conditions on spectral projections and resolvents rather than on the distribution of eigenvalues to handle non–self–adjointness. As Singleton stated in [74], the results obtained for suitable non–self–adjoint operators are of the same order of magnitude.

We will proceed in a different way based on an idea due to Froyland and Dellnitz [25]. They made the fascinating observation that we can associate to every non–reversible Markov chain a reversible Markov chain that possesses the same invariant measure and the same metastable subsets. We will extend this approach for our purpose from the finite state space to the general state space. Consider a Markov process $X_t = \{X_t\}_{t\in\mathbf{T}}$ and assume that its stochastic transition function p is absolutely continuous, hence $p(t,x,\mathrm{d}y) = p(t,x,y)\mu(\mathrm{d}y)$ with density jointly measurable in x and y. For some fixed observation time span $\tau > 0$, consider the discrete–time Markov process $X_n = \{X_{\tau n}\}_{n\in\mathbf{Z}_+}$; its stochastic transition function is given by $p_\tau(n,x,\mathrm{d}y) = p(\tau n,x,\mathrm{d}y)$. We define the **time–reversed Markov process** $Y_n = \{Y_n\}_{n\in\mathbf{Z}_+}$ via its stochastic transition function q_τ given by

$$q_\tau(n,x,\mathrm{d}y) \;=\; q_\tau(n,x,y)\mu(\mathrm{d}y) \;=\; p_\tau(n,y,x)\mu(\mathrm{d}y),$$

which by definition is again absolutely continuous and discrete in time. The map q_τ satisfies the requirements for a stochastic transition function, since $q_\tau(n,x,\mathbf{X}) = 1$ due to invariance of μ, and $\int_{\mathbf{X}} q_\tau(n,x,\mathrm{d}z)q_\tau(1,z,A) = q_\tau(n+1,x,A)$, which implies the Chapman–Kolmogorov equation. If p_τ is reversible then $q_\tau = p_\tau$, as we would expect. The time–reversed Markov process Y_n has two important properties (analogous to [25]):

(i) The probability measure μ is invariant w.r.t. Y_n, since

$$\begin{aligned}\int_{\mathbf{X}} q_\tau(n,x,A)\mu(\mathrm{d}x) &= \int_{\mathbf{X}}\int_A p(\tau n,y,x)\mu(\mathrm{d}y)\mu(\mathrm{d}x)\\ &= \int_A\int_{\mathbf{X}} p(\tau n,y,x)\mu(\mathrm{d}x)\mu(\mathrm{d}y) \;=\; \mu(A).\end{aligned}$$

(ii) The time–reversed Markov process Y_n has the same metastable subsets as the original process X_n, i.e.,

$$p_\tau(n,C,C) \;=\; q_\tau(n,C,C)$$

for arbitrary $C \in \mathcal{A}$ and every $n \in \mathbf{Z}_+$. This is a special case of the general identity $\mu(C)\, p_\tau(n, C, D) = \mu(D)\, q_\tau(n, D, C)$, which is due to

$$\begin{aligned} \mu(C)\, p_\tau(n, C, D) &= \int_C \int_D p_\tau(n, x, y) \mu(\mathrm{d}y) \mu(\mathrm{d}x) \\ &= \int_C \int_D q_\tau(n, y, x) \mu(\mathrm{d}y) \mu(\mathrm{d}x) \\ &= \mu(D)\, q_\tau(n, D, C). \end{aligned}$$

Although neither X_n nor Y_n need to be reversible, we can use them to construct a reversible Markov process (analogous to [25]):

Theorem 3.2 *Define the* ***time–symmetrized Markov process*** $Z_n = \{Z_n\}_{n \in \mathbf{Z}_+}$ *via its* one–step *stochastic transition function*

$$r_\tau(1, x, \mathrm{d}y) = \frac{1}{2} \left[p_\tau(1, x, y) + q_\tau(1, x, y) \right] \mu(\mathrm{d}y).$$

Its n–step version given by the Chapman–Kolmogorov equation. Then, Z_n is invariant w.r.t. to μ, reversible and possesses the same one step metastability as X_n, i.e.,

$$p_\tau(1, C, C) = r_\tau(1, C, C)$$

for arbitrary $C \in \mathcal{A}$.

Proof: The statements about the invariance of μ and the metastability are obvious. Now, let us prove that Z_n is reversible w.r.t. μ, hence we have to prove condition (7). For $A, B \in \mathcal{A}$ we have

$$\begin{aligned} \int_A r_\tau(1, x, B) \mu(\mathrm{d}x) &= \int_A \int_B \frac{1}{2} [p_\tau(1, x, y) + q_\tau(1, x, y)] \mu(\mathrm{d}x) \mu(\mathrm{d}y) \\ &= \int_A \int_B \frac{1}{2} [q_\tau(1, y, x) + p_\tau(1, y, x)] \mu(\mathrm{d}x) \mu(\mathrm{d}y) \\ &= \int_B r_\tau(1, x, A) \mu(\mathrm{d}x); \end{aligned}$$

hence Z_n is reversible. □

If the original Markov process X_t is reversible, then $r_\tau = p_\tau$ and the time-symmetrized Markov process coincides with the original one sampled at rate τ. For the interesting case of an originally non–reversible Markov process this is quite different. In general, we have:

(i) The n–step transition probability $r_\tau(n, \cdot, \cdot)$ is not defined via the sum of the n–step transition probabilities $p_\tau(n, \cdot, \cdot)$ and $q_\tau(n, \cdot, \cdot)$, since already

$$r_\tau(2, x, A) \neq \int_{\mathbf{X}} \frac{1}{2} [\, p_\tau(2, x, A) + q_\tau(2, x, A) \,].$$

This is not surprising, because the left hand side, defined via the Chapman–Kolmogorov equation, involves the product of $p_\tau(1,\cdot,\cdot)$ and $q_\tau(1,\cdot,\cdot)$, while the right hand side does not.

(ii) Due to (i) we have $p_\tau(n,C,C) \neq r_\tau(n,C,C)$ for $n > 1$.

(iii) There exists no continuous–time Markov process $\hat{Z}_t = \{\hat{Z}_t\}_{t\in\mathbf{R}}$ such that Z_n is obtained by sampling $\hat{Z}_t$ at rate τ, i.e., such that $Z_n = \hat{Z}_{n\tau}$ for $n \in \mathbf{Z}_+$. Hence, even if the original Markov process is defined via a stochastic differential equation, this is not the case for the time–symmetrized Markov process.

Sums of transition probabilities are frequently encountered in the Markov chain Monte Carlo theory, where it is well known that a realization of the time–symmetrized Markov process can be performed in two steps: (i) choose randomly one of the two transition functions p_τ or q_τ with equal probability 1/2. (ii) proceed according to the chosen transition function. Repeat this procedure for every discrete time step. However, in order to discretize the propagator corresponding to the time–symmetrized Markov process, we will proceed in a different way, as outlined in Section 5.3.

In view of Theorem 3.2, we conclude that the original, *possibly non–reversible* Markov process X_t possesses a decomposition into metastable subsets (on the timescale τ), if the time–symmetrized Markov process Z_n does. This allows us to apply Theorem 3.1. In particular, the eigenfunctions related to the time–symmetrized Markov process can be used to identify the metastable subsets of the original, non–reversible Markov process. We will exemplify the time–symmetrization approach for the Langevin equation in Section 6.3.