Convergent discretisation schemes for transition path theory for diffusion processes

Nada Cvetković* and Tim Conrad†

Department of Mathematics and Computer Science, Freie Universität Berlin, Arnimallee 6, 14195 Berlin, Germany

Han Cheng Lie‡

Universität Potsdam, Institut für Mathematik, Karl-Liebknecht Str. 24-26, D-14476 Potsdam OT Golm, Germany

Abstract

In the analysis of metastable diffusion processes, Transition Path Theory (TPT) provides a way to quantify the probability of observing a given transition between two disjoint metastable subsets of state space. However, many TPT-based methods for diffusion processes compute the primary objects from TPT, such as the committor and probability current, by solving partial differential equations. The computational performance of these methods is limited by the need for mesh-based computations, the need to estimate the coefficients of the stochastic differential equation that defines the diffusion process, and the use of Markovian processes to approximate the diffusion process. We propose a Monte Carlo method for approximating the primary objects from TPT from sample trajectory data of the diffusion process, without estimating drift or diffusion coefficients. We discretise the state space of the diffusion process using Voronoi tessellations and construct a non-Markovian jump process on the dual Delaunay graph. For the jump process, we define committors, probability currents, and streamlines, and use these to define piecewise constant approximations of the corresponding objects from TPT for diffusion processes. Rigorous error bounds and convergence theorems establish the validity of our approach. A comparison of our method with TPT for Markov chains (Metzner et al., Multiscale Model Simul. 2009) on a triple-well 2D potential provides proof of principle.

1. Introduction

In many applications, one is often interested in understanding the rare transitions of a stochastic dynamical system between two metastable subsets of state space. Here, a 'metastable set' refers to a set in which the system spends a long time on average before exiting. The probability of observing a transition between any two metastable sets tends to be small, i.e. the transitions

---

*Corresponding author: ncvetkovic@math.fu-berlin.de
†conrad@mi.fu-berlin.de
‡hanlie@uni-potsdam.de

1
are rare. In addition, the system exhibits multiscale behaviour, in the sense that the time that the system spends outside the union of metastable sets is significantly smaller compared to the time the system spends inside this union. Such processes occur for example in biology, material sciences, physics and chemistry.

Transition path theory (TPT) is a statistical theory developed for the analysis of transition events of an ergodic diffusion process between two subsets of state space \([0,1]\). Given that metastable Markov processes play an important role as models of complex processes in molecular dynamics, TPT has played an important role in this field and related studies into the dynamics of macromolecules in biological processes; a far from exhaustive list of references in this area of research is \([13, 17, 28, 3, 23, 2, 19, 20, 25, 26]\). In addition to the class of diffusion processes, TPT has been developed for Markov jump processes and Markov chains \([16]\) and has been used to study flows in complex networks \([4]\). TPT has been studied using the tools of stochastic analysis \([12]\) and has been extended to non-ergodic jump processes \([29]\).

We now briefly outline the main concepts and objects of TPT. Let \(S\) denote the state space of the ergodic diffusion process, and let \(A, B\) denote two metastable subsets. By ergodicity, the average behaviour of the system can be obtained from one infinitely long trajectory. The trajectory segments during which the process exits the set \(A\) and enters the set \(B\) before re-entering the set \(A\) are called reactive trajectories. TPT enables the computation of statistics of reactive trajectories, by identifying the vector field that describes the net flow of reactive trajectories in any point of the state space. The integral curves or streamlines of this vector field represent the ‘averaged’ behavior of reactive trajectories of the process, in the sense that for any \(\epsilon > 0\), the \(\epsilon\)-tube about a streamline contains a reactive trajectory. The streamlines are important because they can be used to make predictions about the most likely paths that the diffusion process will take when transitioning from \(A\) to \(B\).

A fundamental object in TPT is the (forward) committor function \(q : S \rightarrow [0,1]\), which can be viewed as a ‘capacitor’ from probability theory \([24]\). For \(x \in S\), \(q(x)\) gives the probability of reaching \(B\) before \(A\), given that the current state of the Markov process is \(x\),

\[
q(x) := \mathbb{P}(X_{\tau_{A\cup B}}(X) \in B | X_0 = x),
\]

where

\[
\tau_{A\cup B}(X) := \inf\{t \geq 0 : X_t \in A \cup B\}
\]

denotes the first hitting time of the process \(\{X_t\}_{t \geq 0}\) with respect to the set \(A \cup B\).

TPT establishes relationships between fundamental objects such as the committor and the parameters that define the Markov process. For example, when the Markov process is a diffusion process, then the committor is the solution to a Dirichlet boundary value problem, defined by the infinitesimal generator of the diffusion. These relationships imply that, when computationally efficient methods for solving PDEs are available – e.g. when the state space is a subset of \(\mathbb{R}^d\) for \(d \leq 3\) – then one can in principle completely bypass the problem of collecting statistics of reactive trajectories. Instead one applies deterministic methods to compute the committor function, the probability current of reactive trajectories and the corresponding streamlines. Using the computed objects, one can describe transitions between \(A\) and \(B\). This can be done efficiently, and can lead to useful visualisations of the behaviour of reactive trajectories, see e.g. \([15]\).

In certain applications, the above-mentioned relationships that TPT establishes are difficult to exploit. This is because complex systems are often high-dimensional, and the curse of dimensionality renders many, if not most, deterministic methods for solving partial differential equations too costly to be practical. Another possible problem is that the parameters are not
known or difficult to estimate. This creates a need for methods that can be applied to complex systems, but that are neither based on solving partial differential equations nor on a priori knowledge of the parameters that define the Markov process. TPT for Markov jump processes avoids solving the partial differential equations, but is limited to discrete processes that are Markovian. Even though TPT for Markov jump processes is used in practice to approximate the dynamics of diffusion processes, e.g., in molecular dynamics, there is to the best of our knowledge no theoretical validation for this approach. In particular, there are no proofs of convergence of the objects of TPT for Markov jump processes to the corresponding objects of TPT for diffusion processes.

1.1. Contributions and outline

The goal of this paper is to present and analyse a method for approximating some key objects of TPT for an ergodic diffusion process that takes values in a continuous, compact state space $S \subset \mathbb{R}^d$ with reflecting boundary conditions. Our method uses a non-Markovian jump process on a discrete state space $I$, where the discrete state space is obtained by Voronoi tessellation of the continuous state space $S$.

Voronoi tessellations are powerful discretisation schemes that have been applied in a wide range of contexts; see the applications mentioned in [1, 6]. In addition, Voronoi tessellations have been used to circumvent the curse of dimensionality in the context of molecular dynamics [11, 27]. The discrete state space $I$ of the non-Markovian jump process is the index set of the Voronoi tessellation of continuous state space $S$.

The TPT objects of the diffusion process that we approximate are the committor, probability current, and streamlines. These objects are continuous functions that either take values in or are defined on the state space $S$ of the diffusion process. To approximate these objects, we construct discrete analogues that either take values in or are defined on the discrete state space $I$ of the jump process. We then use these discrete analogues to construct piecewise constant functions in or on the state space $S$. The present work focuses on finding appropriate definitions of the latter objects and analysing their approximation quality with respect to the continuous counterparts from TPT for diffusion processes.

We highlight some important features of our work. First, the objects we define do not require solving PDEs. Second, the objects we define do not require prior knowledge or estimation of the coefficients of the underlying diffusion process; only trajectory data of the diffusion process is needed. In this sense, our approach is purely data-driven and can be applied to any stochastic dynamical system with continuous paths taking values in $S$. Third, we establish the validity of our method, by proving quantitative error bounds of the approximations that we define for the committor, probability current, and streamlines. We use these bounds to prove convergence of these objects to their counterparts for the underlying ergodic diffusion process, in the ‘continuum limit’, i.e. as the partition becomes infinitely fine. The convergence analysis for larger classes of systems is an interesting question that we do not address here.

The paper proceeds as follows. We describe how the Voronoi tessellation of the continuous state space $S$ leads to a non-Markovian jump process in Section 2. In Section 3 we define the committor, the probability current, and the streamlines associated to the probability current on the discrete state space of the jump process. We use these ‘discrete TPT’ objects to define piecewise constant functions in or on the continuous state space of the diffusion process. For each piecewise constant approximation, we prove a bound on the approximation error with respect to the corresponding ‘continuous TPT object’ of the underlying diffusion process. For each object, the errors are bounded by a constant times the largest diameter over all sets in the
partition. We then use these error bounds to prove convergence as the largest diameter over all sets in the partition decreases to zero; this is the above-mentioned ‘continuum limit’. In Section 4 we present numerical results, in which we compare the performance of our approach with the performance of TPT for Markov chains and TPT for diffusion processes. We conclude in Section 5.

2. Setup

Let $X = \{X_t\}_{t \geq 0}$ be an ergodic diffusion process taking values in a compact subset $S \subset \mathbb{R}^d$. We assume that the boundary of $S$ is sufficiently regular to impose reflecting boundary conditions. Suppose that the invariant measure $\mu$ of $X$ is absolutely continuous with respect to Lebesgue measure with Lebesgue density $m : S \to \mathbb{R}$

$$\mu(A) := \mathbb{P}(X_t \in A) = \int_A m(x)dx, \quad \forall A \in \mathcal{B}(S).$$

A Voronoi tessellation of $S$ associated to a finite set of generators $\{g_1, \ldots, g_n\}$ for some $n \in \mathbb{N}$ is a collection $\{S_1, \ldots, S_n\}$ of nonempty subsets of $S$, where each Voronoi cell is defined by

$$S_i := \{x \in S : |x - g_i|^2 \leq |x - g_j|^2, \quad \forall j \neq i\},$$

where $|\cdot|_2$ denotes the Euclidean distance. That is, $S_i$ is the closed set consisting of all points in state space that are closer in the Euclidean metric to the generator $g_i$ than to any other generator. Since every Voronoi cell $S_i$ is a closed neighbourhood of its generator $g_i$, it has strictly positive Lebesgue measure. Let $\text{int}(A)$ and $\partial A$ denote the interior and boundary of $A \subset S$ respectively. Observe that

$$S = \bigcup_{i=1}^n S_i, \quad S_i \cap S_j = \partial S_i \cap \partial S_j, \quad i \neq j,$$

so that Voronoi cells intersect at most at their boundaries. Consider the following definition.

**Definition 2.1.** Two distinct Voronoi cells $S_i$ and $S_j$ are adjacent if they share a common facet, i.e. if

$$\dim(S_i \cap S_j) = d - 1.$$  

Given a Voronoi tessellation $\{S_i\}_{i \in I}$, $I = \{1, \ldots, n\}$, the dual object is the Delaunay graph $G = (I, E)$ with vertex set $I$ and edge set $E$ consisting of all pairs $(i, j)$ such that $S_i$ and $S_j$ are adjacent.

Recall that, given a nonempty set $A \subset \mathbb{R}^d$, the Euclidean diameter of $A$ is defined by

$$\text{diam}(A) = \sup\{|x - y|_2 : x, y \in A\}.$$  

This leads us to the next definition.

**Definition 2.2.** The width $\rho$ of a Voronoi tessellation $\{S_i\}_{i \in I}$ is the largest Euclidean diameter of the Voronoi cells, i.e.

$$\rho(\{S_i\}_{i \in I}) := \max_{i \in I} \text{diam}(S_i).$$
2.1. Definition of approximating jump process

Let $A$ and $B$ be the metastable sets mentioned earlier. Let $\text{int}(U)$, $\text{cl}(U)$ and $\partial U$ denote the interior, closure, and boundary of an arbitrary set $U$. We make the following assumption.

**Assumption 2.3.** The sets $A$ and $B$ are open and convex sets, such that $\text{cl}(A)$ and $\text{cl}(B)$ are disjoint. There exist disjoint subsets $J, K \subset I$ such that $\text{cl}(A) = \bigcup_{j \in J} S_j$ and $\text{cl}(B) = \bigcup_{k \in K} S_k$.

We justify this assumption as follows. First, when the diffusion process is described by a stochastic differential equation whose drift coefficient is minus the gradient of some potential, then the metastable sets $A$ and $B$ can be taken to be balls centred at two distinct local minima of the potential. Second, it is known that one can approximate any bounded convex set by collections of (bounded) polytopes \cite{22}. Third, Voronoi cells are polytopes, as they are defined by systems of linear inequalities.

Let $\{S_i\}_{i \in I}$ be a Voronoi tessellation with finite index set $I = \{1, \ldots, n\}$, and let $G = (I, E)$ be its associated Delaunay graph. We define a continuous time jump process $Y = \{Y_t\}_{t \geq 0}$ with state space equal to the index set $I$ as follows:

$$Y_t = \begin{cases} i & X_t \in \text{int}(S_i), \\ j & \exists \varepsilon > 0 \text{ s.t. } X_s \in \text{int}(S_j) \forall s \in (t - \varepsilon, t). \end{cases} \quad (2.1)$$

Next, define the first hitting time of $Y$ with respect to the set $J \cup K$:

$$\tau_{J \cup K}(Y) := \inf \{ t \geq 0 : Y_t \in J \cup K \}. \quad (2.2)$$

The second case in (2.1) can be interpreted as follows: if $X_t$ lies on the boundary of a Voronoi cell, then we assign to $Y_t$ the index $j$ where $j$ is the index of the set whose interior contained the trajectory of $Y$ in the ‘most recent past’, i.e. up to but not including the current time $t$. Then, given the first hitting time of $X$ with respect to $A \cup B$ that was defined in (1.2), we have the following result.

**Lemma 2.4.** Suppose that Assumption 2.3 holds, and let $\tau_{A \cup B}(X)$ and $\tau_{J \cup K}(Y)$ be defined as in (1.2) and (2.2) respectively. Then $\tau_{A \cup B}(X) = \tau_{J \cup K}(Y)$.

**Proof.** Fix an arbitrary sample trajectory $X(\omega)$ of $X$, and let $t := \tau_{A \cup B}(X(\omega))$. Given that $A$ and $B$ are open, and given the definition (1.2) of $\tau_{A \cup B}(X)$, it follows that there exist $\varepsilon_1, \varepsilon_2 > 0$ such that $X_s(\omega) \notin A \cup B$ for $s \in (t - \varepsilon_1, t)$, $X_t(\omega) \in \partial A \cup \partial B$, and $X_s(\omega) \in A \cup B$ for $s \in (t, t + \varepsilon_2)$. Suppose first that $X_t(\omega) \in \partial A$ and $X_s(\omega) \in A$ for $s \in (t, t + \varepsilon_2)$. Let $i \in I$ be such that $X_s(\omega) \in \text{int}(S_i)$ for $s \in (t - \varepsilon_1, t)$. By the definition (2.1) of $Y$, it follows that $Y_s(\omega) = i$ for $s \in (t - \varepsilon_1, t)$ and $Y_t(\omega) = i$. Furthermore, given Assumption 2.3, $X_s(\omega) \in A$ for $s \in (t, t + \varepsilon_2)$ implies that $Y_s(\omega) \in J$ for $s \in (t, t + \varepsilon_2)$. If instead $X_t(\omega) \in \partial B$ and $X_s(\omega) \in B$ for $s \in (t, t + \varepsilon_2)$, then $Y_s(\omega) \in K$ for $s \in (t, t + \varepsilon_2)$. Finally, since $t = \tau_{A \cup B}(X(\omega))$, it follows from the definition (2.2) that $\tau_{J \cup K}(Y(\omega)) = \inf\{ t, t + \varepsilon_2 \} = t$. This proves the claim. \qed
3. Definition and convergence analysis of discrete TPT objects

In this section we define transition path theory objects for the jump process $Y$ that was introduced in (2.1). We prove that in the limit of partition width decreasing to zero, the objects we define for the jump process $Y$ converge to the corresponding TPT objects for the diffusion process $X$. More precisely, we define analogues of the committor, probability current, and streamlines for this jump process, and for each such object we prove an error bound in an appropriate metric with respect to its analogue from TPT for diffusion processes.

3.1. Committors

Recall the committor function of the diffusion process $X$ defined in (1.1). For $i \in I$, define

$$
\hat{q}_i = \frac{1}{\mu(S_i)} \langle q, 1_{S_i} \rangle \mu,
$$

(3.1)

where $1_{S_i} : S \rightarrow \{0, 1\}$ is the indicator function of $S_i$ and $\langle \cdot, \cdot \rangle \mu$ is the inner product weighted by the invariant measure $\mu$ of the process $X$, i.e. $\langle v, w \rangle \mu = \int_S v(x) w(x) \mu(dx)$.

Using the collection $\{\hat{q}_i\}_{i \in I}$, we can construct a function that is piecewise constant on the interiors of the Voronoi cells, the projected committor function $\hat{q} : S \rightarrow [0, 1]$:

$$
\hat{q}(x) := \sum_{i \in I} \hat{q}_i 1_{\text{int}(S_i)}(x).
$$

(3.2)

To complete the definition of $\hat{q}$, we need to specify its values on the boundaries of the Voronoi cells. However, since the union of the intersections has Lebesgue measure zero and since we will measure the error of $\hat{q}$ with respect to $q$ in an $L^p$ norm, the values that we prescribe will not be important. One straightforward assignment is as follows: let $\emptyset \neq C \subset I$ be such that $x \in \partial S_c$ for all $c \in C$; then for such $x$, define $\hat{q}(x)$ according to

$$
\hat{q}(x) = \max_{c \in C(x)} \hat{q}_c.
$$

Other assignments are possible, e.g. the minimum or the arithmetic mean of $\hat{q}_c$ over $c \in C$.

Now we define the discrete committor $\tilde{q} : I \rightarrow [0, 1]$, which is the committor that corresponds to the time continuous jump process $\{Y_t\}_{t \geq 0}$ on the discrete state space $I$.

We define the discrete committor by the conditional probability

$$
\tilde{q}_i = \frac{\mathbb{P}(Y_{\tau_{K \cup K}}(Y) \in K, Y_0 = i)}{\mathbb{P}(Y_0 = i)}.
$$

(3.3)

We will prove that $\hat{q}_i = \tilde{q}_i$ for all $i$, using the notion of a regular conditional distribution. To define a regular conditional distribution, we first recall the definition of a stochastic kernel.

**Definition 3.1** (Stochastic kernel). Let $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ be measurable spaces. A map $\kappa : \Omega_1 \times \mathcal{A}_2 \rightarrow [0, \infty]$ is called a stochastic kernel from $\Omega_1$ to $\Omega_2$ if:

(i) $\kappa(\cdot, A_2)$ is $\mathcal{A}_1$-measurable for any $A_2 \in \mathcal{A}_2$, and

(ii) $\kappa(\omega_1, \cdot)$ is a $\sigma$–finite probability measure on $(\Omega_2, \mathcal{A}_2)$ for any $\omega_1 \in \Omega_1$. 

6
Definition 3.2 (Regular conditional distribution). Let \((\Omega_1, A_1, \mathbb{P})\) be a probability space, \((E, \mathcal{E})\) be a measurable space, and \(Y\) be an \(E\)-valued random variable on \((\Omega_1, A_1, \mathbb{P})\). Let \(\mathcal{F} \subset A_1\) be a \(\sigma\)-algebra on \(\Omega_1\). A stochastic kernel \(\kappa_{Y,\mathcal{F}}\) from \((\Omega_1, \mathcal{F})\) to \((E, \mathcal{E})\) is called a regular conditional distribution of \(Y\) given \(\mathcal{F}\) if
\[
\kappa_{Y,\mathcal{F}}(\omega, C) = \mathbb{P}(Y \in C|\mathcal{F})(\omega)
\]
for \(\mathbb{P}\)-almost all \(\omega \in \Omega_1\) and for all \(C \in \mathcal{E}\).

If \(\mathcal{F}\) is generated by a random variable \(X\) defined on \((\Omega_1, A_1, \mathbb{P})\) that takes values in some measurable space \((E', \mathcal{E}')\) then the stochastic kernel \(\kappa_{Y,X}\) from \((E', \mathcal{E}')\) to \((E, \mathcal{E})\) that is defined by
\[
\kappa_{Y,X}(x, C) = \mathbb{P}(Y \in C|X = x) = \kappa_{Y,\sigma(X)}(\omega, C), \quad \forall \omega \in X^{-1}\{x\}
\]
is called a regular conditional distribution of \(Y\) given \(X\).

The existence of the regular conditional distribution of \(Y\) given \(\mathcal{F}\) for a random variable \(Y\) taking values in a Borel space is given in [10, Theorem 8.37].

Proposition 3.3 (Committors are regular conditional probabilities). The committor function \(q\) defined in (1.1) is a regular conditional probability.

Proof. Set the random variables \(X, Y,\) and the set \(C\) in (3.5) to be equal \(X_0, X_{\tau_{A_u,B}(x)},\) and \(B,\) respectively. Then
\[
\kappa_{X_{\tau_{A_u,B}(x)}, X_0}(x, B) = \mathbb{P}(X_{\tau_{A_u,B}(x)} \in B|X_0 \in x) = q(x),
\]
where we used the definition (1.1) of the committor function in the second equation. \(\square\)

The following theorem expresses the conditional expectation of a function of a random variable in terms of regular conditional distribution, and can be found in [10, Theorem 8.38].

Theorem 3.4 (Conditional expectations in terms of regular conditional distributions). Let \(Y\) be a random variable on \((\Omega_1, A_1, \mathbb{P})\) with values in some set \(E,\) and equip \(E\) with the Borel \(\sigma\)-algebra \(\mathcal{E}.\) Let \(\mathcal{F} \subset A_1\) be a \(\sigma\)-algebra and let \(\kappa_{Y,\mathcal{F}}\) be a regular conditional distribution of \(Y\) given \(\mathcal{F} \). Further, let \(f : E \rightarrow \mathbb{R}\) be measurable and \(\mathbb{E}[[f(Y)] < \infty.\) Then
\[
\mathbb{E}[f(Y)|\mathcal{F}](\omega) = \int_E f(y)\kappa_{Y,\mathcal{F}}(\omega, dy)
\]
\(\mathbb{P}\)-almost surely.

We use this theorem to prove the following lemma which expresses the joint probability of a pair of random variables \(X\) and \(Y\) in terms of regular conditional probability of \(Y\) given \(X\).

Lemma 3.5. Let \(X\) and \(Y\) be random variables on \((\Omega_1, A_1, \mathbb{P}),\) taking values in the measurable spaces \((E', \mathcal{E}')\) and \((E, \mathcal{E})\) respectively. Then for any \(C \in \mathcal{E}\) and \(D \in \mathcal{E}',\)
\[
\mathbb{P}(Y \in C, X \in D) = \int_D \kappa_{Y,X}(x, C)(\mathbb{P} \circ X^{-1})(dx).
\]

Proof. By definition of expectation
\[
\mathbb{E}[1_D(X)\mathbb{E}[1_C(Y)|\sigma(X)]] = \int_{\Omega} 1_D(x) \mathbb{E}[1_C(Y)|\sigma(X)](\omega)\mathbb{P}(d\omega).
\]

\[\text{(3.6)}\]
Using Theorem 3.4 in the first equality, (3.5) in the second, and the fact that a stochastic kernel is a probability measure for fixed \( \omega \in \Omega \) in the third equality, we obtain:

\[
\mathbb{E} [1_C(Y) | \sigma(X)] (\omega) = \int_E 1_C(y) \kappa_{Y | \sigma(X)}(\omega, dy) = \int_E 1_C(y) \kappa_{Y,X}(X(\omega), dy) \\
= \kappa_{Y,X}(X(\omega), C).
\] (3.7)

Then,

\[
\int_{\Omega} 1_D(X(\omega)) \mathbb{E}[1_C(Y) | \sigma(X)](\omega) \mathbb{P}(d\omega) = \int_{\Omega} 1_D(X(\omega)) \kappa_{X,Y}(X(\omega), C) \mathbb{P}(d\omega) \\
= \int_E 1_D(x) \kappa_{X,Y}(x, C) \mathbb{P} \circ X^{-1}(dx).
\] (3.8)

where we used (3.7) in the first and the change of variables formula in the second equality.

By using that \( \mathbb{E}[X_1 X_2 | \mathcal{F}] = X_2 \mathbb{E}[X_1 | \mathcal{F}] \) when \( X_2 \) is \( \mathcal{F} \)-measurable in the first equation and the tower law in the second equality we obtain:

\[
\mathbb{E} [1_D(X)|1_C(Y)|\sigma(X)] = \mathbb{E} [\mathbb{E}[1_D(X)1_C(Y)|\sigma(X)]] \\
= \mathbb{E}[1_D(X)1_C(Y)] = \mathbb{P}(Y \in C, X \in D).
\] (3.9)

Finally, combining (3.9), (3.6) and (3.8) proves the claim

\[
\mathbb{P}(Y \in C, X \in D) = \int_E 1_D(x) \kappa_{X,Y}(x, C) \mathbb{P} \circ X^{-1}(dx).
\]

We now use Lemma 3.5 to prove that the projected committor \( \hat{q}_i \) and the discrete committor \( \tilde{q}_i \) are equal for every \( i \in \{1, \ldots, n\} \).

**Proposition 3.6.** Suppose that Assumption 2.3 holds. Let the projected committor \( \hat{q}_i \) and discrete committor \( \tilde{q}_i \) be defined as in (3.1) and (3.3), respectively. Assume that \( X_0 \) is distributed according to the equilibrium distribution \( \mu \). Then \( \hat{q}_i = \tilde{q}_i \), for all \( i \in I \).

**Proof.** Let \( i \in I \) be arbitrary. Recall from (3.1) that

\[
\hat{q}_i = \frac{1}{\mu(S_i)} \langle q, 1_{S_i} \rangle_\mu = \frac{1}{\mu(S_i)} \int_S q(x) 1_{S_i}(x) \mu(dx).
\]

The definition (3.3), the construction of \( Y \), the hypothesis that \( X_0 \) is distributed according to the equilibrium measure \( \mu \), and Lemma 2.4 imply that

\[
\tilde{q}_i = \frac{\mathbb{P}(X_{\tau_{A\cup B}}(x) \in B, X_0 \in S_i)}{\mathbb{P}(Y_0 = i)} \\
= \frac{\mathbb{P}(X_{\tau_{A\cup B}}(x) \in B, X_0 \in S_i)}{\mu(S_i)}.
\]

Thus, to prove the proposition, it suffices to show that

\[
\int_S q(x) 1_{S_i}(x) \mu(dx) = \mathbb{P}(X_{\tau_{A\cup B}}(x) \in B, X_0 \in S_i).
\]
By Proposition 3.3, the left-hand side can be rewritten in terms of a regular conditional probability,

$$\int_S q(x)1_{S_i}(x) \mu(dx) = \int_S 1_{S_i}(x) \kappa_{X_{A\cup B}(x),X_0}(x,B) \mu(dx) = \int_{S_i} \kappa_{X_{A\cup B}(x),X_0}(x,B) \mu(dx).$$

Using that $\mu = P \circ X_0^{-1}$ and Lemma 3.5, we obtain

$$\int_{S_i} \kappa_{X_{A\cup B}(x),X_0}(x,B) \mu(dx) = \int_{S_i} \kappa_{X_{A\cup B}(x),X_0}(x,B) P \circ X_0^{-1}(dx) = P(X_{A\cup B}(X) \in B, X_0 \in S_i),$$

yielding the desired conclusion. \hfill \Box

The following lemma uses the convexity of Voronoi cells and the continuity of the committor function to prove that in every Voronoi cell $S_i$ there exists a point $x_i \in S_i$ at which the continuous committor function $q$ of the diffusion has the value of the projected committor function $\hat{q}_i$. We shall use this lemma later to prove Theorem 3.8.

**Lemma 3.7.** Let $\{S_i\}_{i \in I}$ be a Voronoi tessellation of $S$, and let $\hat{q}_i$ be defined as in (3.1). For every $i \in I$, there exists some $x_i \in S_i$ such that $q(x_i) = \hat{q}_i$.

**Proof.** If $q$ is constant on $S_i$, then it must equal $\hat{q}_i$, so there exist uncountably many $x_i$ that satisfy the desired property. Therefore, suppose that $q$ is not constant on $S_i$, and partition $S_i$ into the disjoint subsets $S_i^- := \{x \in S_i : q(x) < \hat{q}_i\}, S_i^+ := \{x \in S_i : q(x) > \hat{q}_i\}$ and $S_i^0 := \{x \in S_i : q(x) = \hat{q}_i\}$. Since $q$ is continuous and not constant on $S_i$, there must exist some $a \in S_i^-$ and $b \in S_i^+$. It follows from the intermediate value theorem that there exists a $t \in (0,1)$ such that $x_i(t) := (1-t)a + tb$ satisfies $q(x_i(t)) = \hat{q}_i$. Since $a,b \in S_i$ and since any Voronoi cell $S_i$ is convex, it follows that $x_i(t)$ belongs to $S_i$. \hfill \Box

Note that the assumption of convexity in Lemma 3.7 is not necessary, provided that each $S_i$ is pathwise connected. Next, we prove an error bound for the error incurred when we approximate the true committor $q$ with the projected committor $\hat{q}$ defined in (3.2).

**Theorem 3.8 (Error bound for projected committor).** Suppose that the committor $q : S \to [0,1]$ has bounded derivatives of first order, i.e. $\nabla q \in L^\infty$, and let $p \in [1,\infty)$. Furthermore, assume there exists $K > 0$ independent of $x, x_i$, and $i$, such that

$$|q(x) - q(x_i) - (\nabla q(x_i), x - x_i)| \leq K |x - x_i|_2, \quad (3.10)$$

for $|x - x_i| \leq \rho$ and $\rho < 1$. Then there exists some $C = C(q) > 0$, such that for any Voronoi tessellation $\{S_i\}_{i \in I}$ of $S$ with width $\rho < 1$, the corresponding projected committor function $\hat{q}$ satisfies

$$\|q - \hat{q}\|_{L^p(\mu)} \leq C \rho.$$

In particular, as the width of the Voronoi tessellation decreases to zero, the $L^p(\mu)$ error of $\hat{q}$ decreases at least linearly with $\rho$. 

9
Proof. It suffices to prove the first statement, since the second statement follows from the first. Fix an arbitrary \( p \in [1, \infty) \), and fix an arbitrary \( i \in I \). Let the partition width \( \rho \) of \( \{S_i\}_{i \in I} \), be \( \rho < 1 \). By Lemma 3.7, there exists an \( x_i \in S_i \) such that \( q(x_i) = \hat{q}_i = q|_{S_i} \). Computing the \( L^p(\mu) \)-error of the restrictions of \( q \) and \( \hat{q} \) to \( S_i \), and using that \( \hat{q}|_{S_i} = q(x_i) \) by definition of \( x_i \), we obtain

\[
\| (q - \hat{q})|_{S_i} \|_{L^p(\mu)} = \int_{S_i} |q(x) - \hat{q}(x)|^p \mu(dx) = \int_{S_i} |q(x) - q(x_i)|^p \mu(dx)
\]

\[
= \int_{S_i} |q(x) - q(x_i) - \langle \nabla q(x_i), x - x_i \rangle + \langle \nabla q(x_i), x - x_i \rangle|^p \mu(dx)
\]

\[
\leq 2^{p-1} \left( \int_{S_i} (K|x - x_i|_2^2)^p \mu(dx) + \int_{S_i} |\nabla q(x_i), x - x_i|^p \mu(dx) \right)
\]

\[
\leq 2^{p-1} \left( \int_{S_i} (K|x - x_i|_2^2)^p \mu(dx) + \int_{S_i} (\|\nabla q\|_{\infty} |x - x_i|^2)^p \mu(dx) \right)
\]

\[
\leq 2^{p-1} \rho^p \mu(S_i) (K^p + \|\nabla q\|_{\infty}^p)
\]

where we used the inequality \((a + b)^p \leq 2^{p-1}(a^p + b^p)\) in the first inequality, (3.10) and the Cauchy-Schwarz inequality in the second and the fact that \( x, x_i \in S_i \) implies that \( |x - x_i| \leq \rho \) in the third. Therefore, we have that

\[
\|q - \hat{q}\|_{L^p(\mu)} = \sum_{i \in I} \|q - \hat{q}|_{S_i}\|_{L^p(\mu)} \leq 2^{p-1} \rho^p (K^p + \|\nabla q\|_{\infty}^p) \sum_{i \in I} \mu(S_i) \leq (2\rho(\|\nabla q\|_{\infty} + K))^p
\]

where we used that \( \sum_{i \in I} \mu(S_i) = 1 \), \( 2^{p-1} < 2^p \) and \( a^p + b^p \leq (a + b)^p \), for \( a, b \geq 0 \) and \( p \geq 1 \). This proves the claim for \( C = C(q) = 2(\|\nabla q\|_{\infty} + K) \).

In an analogous way to how we defined the function \( \hat{q} : S \rightarrow [0, 1] \) using the finite collection \( \{\hat{q}_i\}_{i \in I} \) of values, we can define the discrete committer function \( \tilde{q} : S \rightarrow [0, 1] \) using the collection \( \{\tilde{q}_i\}_{i \in I} \). This yields the following corollary.

Corollary 3.9 (Error bound for discrete committer). Suppose that Assumption 2.3 and the assumptions of Theorem 3.8 holds. Then for the same scalar \( C \), it holds that for any Voronoi tessellation \( \{S_i\}_{i \in I} \) of \( S \) with width \( \rho \), the function \( \tilde{q} \) satisfies

\[
\|q - \tilde{q}\|_{L^p(\mu)} \leq C\rho,
\]

and the \( L^p(\mu) \) error of \( \tilde{q} \) decreases linearly with the width \( \rho \).

Proof. The result follows from Theorem 3.8 and Proposition 3.6.

3.2. Probability current

In this section we define a discrete probability current \( \tilde{J}_{AB} \) which is obtained by observing the jump process \( Y \) in the discrete state space. As with the preceding sections, we consider a Voronoi tessellation \( \{S_i\}_{i \in I} \), and we assume that we can observe the time continuous process \( \{Y_t\}_{t \geq 0} \), in the sense that we can detect which Voronoi cell \( S_i \) contains \( X_t \) for any \( t \geq 0 \). However, we cannot detect the exact location of \( X_t \). In this section we will use the fact that each Voronoi cell is a bounded, convex, \( d \)-dimensional polytope.

In TPT for diffusion processes, the probability current is a function \( J_{AB} : S \setminus (A \cup B) \rightarrow \mathbb{R}^d \) such that at any point \( x \in S \setminus (A \cup B) \), \( J_{AB}(x) \) represents the net flux of reactive trajectories
from $A$ to $B$ through that point. For any $C \subset S \setminus (A \cup B)$ with $(d-1)$-dimensional boundary $\partial C$, $J_{AB}$ is defined implicitly via

$$
\lim_{s \to 0^+} \lim_{T \to \infty} \frac{1}{s} \int_{\mathbb{R} \cap [0,T]} 1_C(X_t) 1_{C^c}(X_{t+s}) - 1_{C^c}(X_t) 1_C(X_{t+s}) \, dt = \int_{\partial C} J_{AB}(y) \cdot n_C(y) \, d\sigma_C(y)
$$

where $n_C(y)$ denotes the outward facing unit normal to $C$ at some point $y \in \partial C$, $d\sigma_C$ denotes the surface measure on $C$ and $\mathbb{R}$ stands for reactive times, i.e. times during which the trajectory of the process $X$ is reactive [14]. Notice that for a Voronoi cell $S_i$ we have

$$
\int_{\partial S_i} J_{AB}(y) \cdot n_{S_i}(y) \, d\sigma_{S_i}(y) = \sum_{k \in \mathcal{N}_i} \int_{\partial S_i \cap \partial S_k} J_{AB}(y) \cdot n_{ik}(y) \, d\sigma_{\partial S_i \cap \partial S_k}(y), \quad (3.11)
$$

where $\mathcal{N}_i$ denotes the set of indices of cells adjacent to $S_i$, i.e.

$$
\mathcal{N}_i := \{ j \in I : \dim(\partial S_i \cap \partial S_j) = d - 1 \},
$$

and $n_{ik}$ is the unit vector that points out of $S_i$ and is orthogonal to the hyperplane that contains the facet $\partial S_i \cap \partial S_k$. Thus, for any $k \in \mathcal{N}_i$ we have

$$
\alpha_{ik} := \lim_{s \to 0^+} \lim_{T \to \infty} \frac{1}{s} \int_{\mathbb{R} \cap [0,T]} 1_{S_i}(X_t) 1_{S_k}(X_{t+s}) - 1_{S_k}(X_t) 1_{S_i}(X_{t+s}) \, dt = \int_{\partial S_i \cap \partial S_k} J_{AB}(y) \cdot n_{ik} \, d\sigma_{\partial S_i \cap \partial S_k}(y). \quad (3.12)
$$

To define the discrete probability current $\tilde{J}_{AB} : S \setminus (A \cup B) \to \mathbb{R}^d$, we shall assume that the discrete probability current is piecewise constant on the interiors of the Voronoi cells. In particular, for each cell $S_i$, we shall assume that $\tilde{J}_{AB}$ equals some vector $\tilde{J}_{AB,i} \in \mathbb{R}^d$ on $\text{int}(S_i)$. In order to compute $\tilde{J}_{AB,i}$ for all $i \in I$, we will make an additional assumption that $\tilde{J}_{AB}$ is constant and equal to $\tilde{J}_{AB,i}$ on $S_i$, not just the interior of $S_i$. The latter assumption will incur an approximation error in the $\tilde{J}_{AB,i}$ vectors. In Theorem 3.11, we shall control this error under the hypothesis of Lipschitz continuity of the continuous probability current $J_{AB}$, and in Section 3.3, we will describe a recursive procedure for defining the discrete probability current $\tilde{J}_{AB}$ on the boundaries of the Voronoi cells. For this section, however, we will not consider the values of the discrete probability current on the boundaries of the Voronoi cells, beyond requiring that they result in a well-defined function $\tilde{J}_{AB}$ on $S \setminus (A \cup B)$ that is piecewise constant on the interiors of the Voronoi cells.

Let $i \in I$ be the index of a Voronoi cell $S_i$ in the partition. Under the assumption that $\tilde{J}_{AB}$ is constant on $S_i$, we obtain the relation

$$
\int_{\partial S_i \cap \partial S_k} J_{AB}(y) \cdot n_{ik} \, d\sigma_{\partial S_i \cap \partial S_k}(y) = \tilde{J}_{AB,i} \cdot n_{ik} \sigma(\partial S_i \cap \partial S_k), \quad \forall k \in \mathcal{N}_i.
$$

Since we can observe the process $\{Y_t\}_{t \geq 0}$, we can approximate the quantity $\alpha_{ik}$ by using sample data from reactive trajectories. Combining the preceding equation with (3.12) yields

$$
\alpha_{ik} = n_{ik} \cdot \tilde{J}_{AB,i} \sigma(\partial S_i \cap \partial S_k), \quad \forall k \in \mathcal{N}_i.
$$

11
Let \( \hat{\alpha}_{ik} := \alpha_{ik} / \sigma(\partial S_i \cap \partial S_k) \), let \( \hat{\alpha}_i \in \mathbb{R}^{|N_i|} \) denote the vector with \( \hat{\alpha}_{ik} \) as components, and let \( N_i \in \mathbb{R}^{|N_i| \times d} \) denote the matrix whose \( k \)-th row is given by \( n_{ik}^\top \). Then we may rewrite the previous equation as the matrix-vector equation
\[
\hat{\alpha}_i = N_i \tilde{J}_{AB,i} \in \mathbb{R}^{|N_i|}.
\]  (3.13)

For (3.13) to admit a solution \( \tilde{J}_{AB,i} \), we need that \( \hat{\alpha}_i \) belongs to the column space of \( N_i \). However, this need not hold in general, since Proposition A.1 indicates that \( |N_i| \geq d + 1 \), so the column space of \( N_i \) will be a subspace of strictly positive codimension. To solve this problem, we first establish the following important fact about the matrix \( N_i \).

**Lemma 3.10** (\( N_i \) has full rank). Let \( \{S_i\}_{i \in I} \) be a Voronoi tessellation of \( S \), let \( i \in I \) be arbitrary, and let \( n_{ik} \in \mathbb{R}^d \) be the unit normal on \( \partial S_i \cap \partial S_k \) exterior to \( S_i \). Then the corresponding matrix \( N_i \in \mathbb{R}^{(|N_i| \times d)} \) with \( k \)-th row equal to \( n_{ik}^\top \) has rank \( d \).

**Proof.** Since every Voronoi cell \( S_i \) is a \( d \)-dimensional polytope in \( \mathbb{R}^d \), Proposition A.1 implies that \( S_i \) has at least \( d + 1 \) facets. Thus \( |N_i| \) is at least \( d + 1 \). Since \( S_i \) has at least \( d + 1 \) facets, it has at least \( d + 1 \) corresponding outer unit normals. By Corollary A.2, there exist \( d \) linearly independent outer unit normals of \( S_i \), which proves the claim. \( \square \)

As a consequence of the lemma above, we can substitute the problem given in equation (3.13) with the normal equations
\[
M_i \tilde{J}_{AB,i} := (N_i^\top N_i) \tilde{J}_{AB,i} = N_i^\top \hat{\alpha}_i =: \beta_i \in \mathbb{R}^d.
\]  (3.14)

Let \( \sigma_{\min}(\cdot) \) and \( \sigma_{\max}(\cdot) \) denote real-valued functions on matrices that yield the smallest and largest singular value respectively. For each Voronoi cell \( S_i \), we define its smallest and largest singular values by the smallest and largest singular values \( \sigma_{\min}(N_i) \) and \( \sigma_{\max}(N_i) \) of the matrix \( N_i \):
\[
\sigma_{\min}(S_i) := \sigma_{\min}(N_i) = \sqrt{\sigma_{\min}(M_i)}, \quad \sigma_{\max}(S_i) := \sigma_{\max}(N_i) = \sqrt{\sigma_{\max}(M_i)}.
\]  (3.15)

Since \( N_i \) has full rank by Lemma 3.10, it follows that \( M_i := N_i^\top N_i \) has full rank, so there exists a unique solution \( \tilde{J}_{AB,i} \) to (3.14). Solving (3.14) for every \( i \in I \) yields the collection \( \{J_{AB,i}\}_{i \in I} \) corresponding to the partition \( \{S_i\}_{i \in I} \). We then define the discrete probability current \( \tilde{J}_{AB} : S \setminus (A \cup B) \rightarrow \mathbb{R}^d \) so that \( \tilde{J}_{AB}(x) = \tilde{J}_{AB,i} \) for all \( x \in \text{int}(S_i) \). For the purposes of establishing an error bound in the \( L^2(\mu) \) metric, the values of the discrete probability current \( \tilde{J}_{AB} \) on the facets is not important, because the \( \mu \)-measure of the facets is zero. This leads to the following theorem.

**Theorem 3.11** (Error bound for discrete probability current). Let \( \{S_i\}_{i \in I} \) be a Voronoi tessellation of \( S \) with width \( \rho \). Assume that the continuous probability current \( J_{AB} : S \setminus (A \cup B) \rightarrow \mathbb{R}^d \) of the diffusion process is globally Lipschitz with Lipschitz constant \( L \). Then the discrete probability current \( \tilde{J}_{AB} : S \setminus (A \cup B) \rightarrow \mathbb{R}^d \) is such that for every \( i \in I \setminus (J \cup K) \),
\[
\left| \tilde{J}_{AB,i} - J_{AB}(x) \right|_2 \leq \rho \frac{(d|N_i|)^{1/2}L}{\sigma_{\min}(S_i)}, \quad \forall x \in \text{int}(S_i).
\]  (3.16)

In particular, we have
\[
\left\| \tilde{J}_{AB} - J_{AB} \right\|_{L^2(S \setminus (A \cup B), \mu; \mathbb{R}^d)} \leq \rho \max_{i \in I \setminus (J \cup K)} \frac{(d|N_i|)^{1/2}L}{\sigma_{\min}(S_i)}.
\]  (3.17)
Proof. Since the union of the facets in any tessellation has \( \mu \)-measure zero and since the \( \{ \text{int}(S_i) \}_{i \in I} \) are disjoint, it follows that

\[
\| \tilde{J}_{AB} - J_{AB} \|_{L^2(S \setminus (A \cup B), \mu; \mathbb{R}^d)}^2 = \sum_{i \in I \setminus (J \cup J')} \int_{\text{int}(S_i)} \| \tilde{J}_{AB,i} - J_{AB}(x) \|_{L^2}^2 \, dx.
\]

Combined with the fact that \( \mu(S \setminus (A \cup B)) \leq 1 \), it follows that to prove (3.17), we only need to prove (3.16). Fix an arbitrary \( x \in S_i \). Since \( M_i \) is invertible by Lemma 3.10, we may use (3.15) to obtain

\[
\| \tilde{J}_{AB,i} - J_{AB}(x) \|_2^2 = \left| M_i^{-1} M_i \left( \tilde{J}_{AB,i} - J_{AB}(x) \right) \right|_2^2 \\
\leq \sigma_{\text{max}}(M_i^{-1}) \left| M_i \left( \tilde{J}_{AB,i} - J_{AB}(x) \right) \right|_2 \\
= \frac{1}{\sigma_{\text{min}}(M_i)} \left| M_i \left( \tilde{J}_{AB,i} - J_{AB}(x) \right) \right|_2 \\
= \frac{1}{\sigma_{\text{min}}(M_i)} \left| M_i \left( \tilde{J}_{AB,i} - J_{AB}(x) \right) \right|_2.
\]

Let \( G \in \mathbb{R}^{d_1 \times d_2} \), \( g_i^\top \) be the \( i \)-th row of \( G \), and \( v \in \mathbb{R}^{d_2} \); then

\[
\| Gv \|_2^2 = \sum_{i=1}^{d_1} |g_i^\top v|^2 \leq \sum_{i=1}^{d_1} |g_i|^2 \| v \|_2^2.
\]

We will want to use the above inequality with \( G = N_i^\top \) and \( v = N_i(\tilde{J}_{AB,i} - J_{AB}(x)) \). Recall that the rows of \( N_i \) are outer unit normals of the facets of the Voronoi cell \( S_i \). Thus the rows of \( N_i^\top \) are the columns of \( N_i \), which have squared Euclidean norm of at most \( \#N_i \), since the absolute value of each entry of \( N_i \) is at most 1. This implies

\[
\left| N_i \left( \tilde{J}_{AB,i} - J_{AB}(x) \right) \right|_2^2 = \left| N_i^\top N_i \left( \tilde{J}_{AB,i} - J_{AB}(x) \right) \right|_2^2 \leq \sum_{i=1}^{d_1} \#N_i \left| N_i \left( \tilde{J}_{AB,i} - J_{AB}(x) \right) \right|_2^2.
\]

Next, from (3.13) we obtain

\[
\left| N_i(\tilde{J}_{AB,i} - J_{AB}(x)) \right|_2^2 = \sum_{j=1}^{d} |n_{ij} \cdot (\tilde{J}_{AB,i} - J_{AB}(x))|^2 = \sum_{j=1}^{d} |\alpha_{ij} - n_{ij} \cdot J_{AB}(x)|^2.
\]

Using (3.12), the Cauchy-Schwarz inequality, the fact that \( |n_{ij}|_2 = 1 \) for all \( i \in I \) and \( j \in N_i \), the Lipschitz continuity of \( J_{AB} \), and the fact that \( |x - y| \leq \rho \) for any \( x, y \in \partial S_i \subset S_i \), we have

\[
|\alpha_{ij} - n_{ij} \cdot J_{AB}(x)| = \left| \frac{1}{\sigma(\partial S_i \cap \partial S_j)} \int_{\partial S_i \cap \partial S_j} n_{ij} \cdot (J_{AB}(y) - J_{AB}(x)) \, d\sigma_{\partial S_i \cap \partial S_j}(y) \right| \\
\leq \frac{1}{\sigma(\partial S_i \cap \partial S_j)} \int_{\partial S_i \cap \partial S_j} |n_{ij}|_2 |J_{AB}(y) - J_{AB}(x)|_2 \, d\sigma_{\partial S_i \cap \partial S_j}(y) \\
\leq \frac{1}{\sigma(\partial S_i \cap \partial S_j)} \int_{\partial S_i \cap \partial S_j} L \left| y - x \right|_2 \, d\sigma_{\partial S_i \cap \partial S_j}(y) \\
\leq \frac{L \rho}{\sigma(\partial S_i \cap \partial S_j)} \sigma(\partial S_i \cap \partial S_j) = L \rho.
\]

Combining the preceding inequalities yields (3.16), as desired. ■
Define the smallest singular value of a Voronoi tessellation \( \{S_i\}_{i \in I} \) via
\[
\sigma_{\text{min}}(\{S_i\}_{i \in I}) := \min_{i \in I} \sigma_{\text{min}}(S_i),
\]
and let
\[
\#N_{\text{max}}(\{S_i\}_{i \in I}) := \max_{i \in I}(\#N_i)
\]
The following result is a corollary of Theorem 3.11.

**Corollary 3.12.** Let \( \{\rho_k\}_{k \in \mathbb{N}} \subset (0, \infty) \) be a sequence decreasing to zero. Let \( \{\{S^k_i\}_{i \in I(k)}\}_{k \in \mathbb{N}} \) be a sequence of Voronoi tessellations of \( S \setminus (A \cup B) \) with index set \( I(k) \) and width \( \rho_k \), such that there exist finite \( K_1 \in \mathbb{N} \) and \( K_2 > 0 \) that do not depend on \( k \in \mathbb{N} \) and that satisfy
\[
\max_{i \in I(k)} \#N_i \leq K_1, \quad (3.18)
\]
\[
\sigma_{\text{min}}(\{S^k_i\}_{i \in I(k)}) \geq K_2. \quad (3.19)
\]
Let \( \tilde{J}_{AB}^k : S \setminus (A \cup B) \to \mathbb{R}^d \) be the discrete probability current corresponding to \( \{S^k_i\}_{i \in I(k)} \). If the continuous probability current \( J_{AB} \) of the diffusion process is globally Lipschitz with constant \( L \), then for all \( k \in \mathbb{N} \),
\[
\left| \tilde{J}_{AB}^k(x) - J_{AB}(x) \right|_2 \leq \rho_k \frac{(dK_1)^{1/2}L}{K_2^2}, \quad \forall x \in \text{int}(S_i), \forall i \in I(k), \quad (3.20)
\]
and
\[
\left\| \tilde{J}_{AB}^k - J_{AB} \right\|_{L^2(S \setminus (A \cup B), \mu; \mathbb{R}^d)} \leq \rho_k \frac{(dK_1)^{1/2}L}{K_2^2}. \quad (3.21)
\]

**Proof.** Since the right-hand side of (3.20) does not depend on \( x \), squaring both sides of the inequality and integrating over \( S \setminus (A \cup B) \) with respect to \( \mu \) yields (3.21). Thus it suffices to observe that (3.20) follows from (3.16), (3.18), and (3.19). \( \square \)

**Remark 3.13.** There exist Voronoi tessellations that satisfy both conditions (3.18) and (3.19) of Corollary 3.12. For example, the condition (3.18) is satisfied for Voronoi tessellations where each Voronoi cell is a d-simplex, in which case \( K_1 = d + 1 \). In fact, \( K_1 = d + 1 \) is the smallest upper bound possible; see Proposition A.1. On the other hand, (3.19) is satisfied when the partition sets are homothetic images of a single bounded convex polytope, for example; this follows from a result concerning the so-called degeneracy ratio of bounded convex polytopes [5].

### 3.3. Streamlines

We now define discrete streamlines of jump process \( Y \), using the discrete probability current defined in the previous section, and prove that the discrete streamlines converge to the streamlines of the diffusion process \( X \). In TPT for diffusion processes, a streamline between the reactant set \( A \) and the product set \( B \) for a given initial condition \( s_0 \in \partial A \) is an integral curve of the probability current starting at \( s_0 \), i.e. a streamline is the solution \( \{s(t)\}_{t \in [0,T(s_0)]} \) of the initial value problem
\[
s(0) = s_0, \quad \frac{ds}{dt}(t) = J_{AB}(s(t)), \quad t \in [0,T(s_0)] \quad (3.22a)
\]
\[
T(s_0) := \inf\{t > 0 : s(t) \in B\}. \quad (3.22b)
\]
Equivalently, we may define a streamline according to
\[
s(t) = s_0 + \int_0^t J_{AB}(s(r)) \, dr, \quad 0 \leq t \leq T(s_0).
\] (3.23)

Throughout this section, we will make the following assumption.

**Assumption 3.14.** For all \( s_0 \in \partial A \), \( T(s_0) \) is finite and strictly positive.

Recall that in Section 3.2, we defined the discrete probability current associated to a Voronoi tessellation as being a vector field that is piecewise constant on the interiors of the Voronoi cells. For the purposes of obtaining error bounds with respect to the \( L^2(S \setminus (A \cup B), \mu; \mathbb{R}^d) \) metric, the values on the set of boundary points of all the cells in the tessellation were not important, since the set of boundary points has \( \mu \)-measure zero. We shall define the discrete streamline with initial condition \( \tilde{s}_0 \) as the integral curve of the discrete probability current \( J_{AB} \) starting at \( \tilde{s}_0 \):
\[
\tilde{s}(t) = \tilde{s}_0 + \int_0^t \tilde{J}_{AB}(\tilde{s}(r)) \, dr, \quad 0 \leq t \leq T(\tilde{s}_0).
\] (3.24)

These streamlines will necessarily intersect the boundaries of the Voronoi cells. Thus we need to find reasonable definitions of the discrete probability current on the boundaries of cells. To do this, recall that the discrete probability current is constant on the interiors of the Voronoi cells. Let \( \tilde{s}_0 \in \partial A \). By Assumption 2.3, it follows that there exists some \( i \in I \setminus (J \cup K) \) such that \( \tilde{s}_0 \in \partial S_i \). Let the vector field at \( \tilde{s}_0 \) to be \( J_{AB,i} \) where
\[
i^*_i := \text{argmax} \left\{ \left| \tilde{J}_{AB,i} \right|_2 : i \in I(\tilde{s}_0) \right\}, \quad I(\tilde{s}_0) := \{ i \in I \setminus (J \cup K) : \tilde{s}_0 \in \partial S_i \}.
\]

Above, the set \( I(\tilde{s}_0) \) is the set of indices of Voronoi cells whose boundaries contain \( \tilde{s}_0 \). From this set we choose the index \( i^*_i \), and define the discrete streamline so that it is an integral curve of the constant vector field that is equal to \( J_{AB,i} \). Note that if the condition (3.18) holds, then \( \# I(\tilde{s}_0) \leq K_1 \). Since \( J_{AB,i} \) is constant and \( S_i \) is compact, there exists \( \tau_1 > 0 \) such that \( \tilde{s}(\tau_1) \in \partial S_i \setminus \{ \tilde{s}_0 \} \), and \( \{ \tilde{s}(t) : 0 \leq t \leq \tau_1 \} \) is a curve in \( S_i \) with constant velocity vector (for \( 0 < t < \tau_1 \)) equal to \( J_{AB,i} \). Recursive application of this algorithm is summarised as
\[
i^*_{k+1} := \text{argmax} \left\{ \left| \tilde{J}_{AB,i} \right|_2 : i \in I(\tilde{s}(\tau_k)) \setminus \{ i^*_k \} \right\} \quad (3.25a)
\]
\[
I(\tilde{s}(\tau_k)) := \{ i \in I \setminus (J \cup K) : \tilde{s}(\tau_k) \in \partial S_i \} \quad (3.25b)
\]
\[
\tau_{k+1} := \inf \{ t > \tau_k : \tilde{s}(\tau_k) + (t - \tau_k)\tilde{J}_{AB,i^*_{k+1}} \notin \text{int}(S_{i^*_{k+1}}) \}, \quad (3.25c)
\]

where the recursion terminates when \( \tilde{s}(\tau_K) \in \partial S_k \) for \( k \in K \), i.e. when the discrete streamline reaches \( \partial B \).

If we set \( \tau_0 = 0 \), then (3.25) yields a collection \( \{ [\tilde{s}(\tau_k) \tilde{s}(\tau_{k+1})] \}_{k=0,1,...} \) of straight line segments \( [\tilde{s}(\tau_k) \tilde{s}(\tau_{k+1})] = \{ \tilde{s}(t) : \tau_k \leq t \leq \tau_{k+1} \} \), where each segment has constant velocity vector for \( \tau_k < t < \tau_{k+1} \) equal to \( J_{AB,i^*_{k}} \) and \( \tau_k, \tau_{k+1} \in \partial S_{i^*_{k+1}} \) are distinct.

**Remark 3.15.** The recursion (3.25) ensures that the discrete probability current at some boundary point \( z \) is the discrete probability current of one of the Voronoi cells that contain \( z \).

We now apply Corollary 3.12 to prove convergence of the discrete streamline starting at some point \( s_0 \) to the continuous streamline that starts at \( s_0 \).
Theorem 3.16 (Error bound for discrete streamlines). Let \( \{ \rho_k \}_{k \in \mathbb{N}} \) and \( \{ \{ S^k_i \}_{i \in I(k)} \}_{k \in \mathbb{N}} \) be a corresponding family of partitions. Fix \( s_0 \in \partial A \), and let \( s \) denote the continuous streamline defined by (3.23) with \( s(0) = s_0 \) and let \( \{ \tilde{s}^k \}_{k \in \mathbb{N}} \) denote the sequence of discrete streamlines corresponding to \( \{ \{ S^k_i \}_{i \in I(k)} \}_{k \in \mathbb{N}} \) such that (3.24) and (3.25) hold, and \( \tilde{s}^k(0) = s_0 \) for all \( k \in \mathbb{N} \). If the continuous probability current \( J_{AB} \) is Lipschitz continuous with constant \( L \) and if \( \{ S^k_i \}_{i \in I(k)} \) satisfy (3.18) and (3.19) for some finite \( K_1 \in \mathbb{N} \) and \( K_2 > 0 \) that do not depend on \( k \), then

\[
\left\| \tilde{s}^k - s \right\|_{L^2([0,T(s_0)],dt;\mathbb{R}^d)} \leq C \rho_k, \quad \forall k \in \mathbb{N},
\]

where \( C \) does not depend on \( k \).

Proof. Let \( k \in \mathbb{N} \), and \( \{ S^k_i \}_{i \in I(k)} \) be a partition of width \( \rho_k \). Let \( \tilde{J}^k_{AB} \) be the discrete probability current corresponding to \( \{ S^k_i \}_{i \in I(k)} \), and let \( \tilde{s}^k \) be the discrete streamline constructed using (3.24), (3.25), \( \tilde{J}^k_{AB} \), and the initial condition \( \tilde{s}^k(0) = s_0 \). It follows from (3.23) that

\[
\left| s(t) - \tilde{s}^k(t) \right|_2 \leq \int_0^t \left| J_{AB}(s(r)) - \tilde{J}^k_{AB}(\tilde{s}^k(r)) \right|_2 dr.
\]

By the triangle inequality and Lipschitz continuity of \( J_{AB} \), we have

\[
\left| J_{AB}(s(r)) - \tilde{J}^k_{AB}(\tilde{s}^k(r)) \right|_2 \leq \left| J_{AB}(s(r)) - J_{AB}(\tilde{s}^k(r)) \right|_2 + \left| J_{AB}(\tilde{s}^k(r)) - \tilde{J}^k_{AB}(\tilde{s}^k(r)) \right|_2.
\]

for all \( 0 \leq r \leq T(s_0) \). Recall that (3.20) of Corollary 3.12 yields

\[
\left| \tilde{J}^k_{AB}(x) - J_{AB}(x) \right|_2 \leq \rho_k \frac{(dK_1)^{1/2} L}{K_2^2}, \quad \forall x \in \text{int}(S^k), \forall i \in I(k).
\]

Given the recursion (3.25), it follows from Remark 3.15 that the above inequality can be extended to apply to \( x \) that lie on the boundaries of one or more Voronoi cells. This is because according to (3.25), the probability current vector that we assign to a boundary point \( x \) is the probability current vector that applies in the interior of one of the Voronoi cells that contains \( x \). In other words, \( \tilde{J}^k_{AB}(x) = \tilde{J}^k_{AB,i} \), for one of the \( i \) such that \( x \in \partial S_i \). Combining the preceding estimates yields

\[
\left| s(t) - \tilde{s}^k(t) \right|_2 \leq \int_0^t \left( \rho_k \frac{(dK_1)^{1/2} L}{K_2^2} + L \left| s(r) - \tilde{s}^k(r) \right|_2 \right) dr
\]

\[
\leq \rho_k \frac{(dK_1)^{1/2} L}{K_2^2} t + L \int_0^t \left| s(r) - \tilde{s}^k(r) \right|_2 dr
\]

By the Gronwall-Bellman inequality, it follows that

\[
\left| s(t) - \tilde{s}^k(t) \right|_2 \leq \rho_k \frac{(dK_1)^{1/2} L}{K_2^2} \exp \left( LT(s_0) \right) T(s_0), \quad \forall t \in [0,T(s_0)].
\]

We therefore obtain

\[
\left( \int_0^{T(s_0)} \left| \tilde{s}^k(r) - s(r) \right|_2^2 dr \right)^{1/2} \leq \rho_k \frac{(dK_1)^{1/2} L}{K_2^2} \exp \left( LT(s_0) \right) T(s_0)^{3/2},
\]

which proves the claim. \( \square \)
4. Numerical results

In this section we implement our method on a well-known example. We compare the results of our method with the method of transition path theory for Markov chains [14, 16], as well as the transition path theory for diffusion processes [8].

The example we choose is a 2-dimensional Smoluchowski diffusion process given by the following stochastic differential equation:

\[ \text{d}X_t = -\Gamma^{-1}\nabla V(X_t)\text{d}t + \sqrt{2\beta^{-1}\Gamma^{-\frac{1}{2}}}\text{d}W_t, \]

where \( V : \mathbb{R}^d \to \mathbb{R} \) denotes the potential function, \( \beta \) is an inverse temperature parameter, i.e. \( \beta = (k_B T)^{-1} \) with Boltzmann constant denoted by \( k_B \) and temperature \( T \), and \( \Gamma \in \mathbb{R}^{d \times d} \) is a diagonal matrix with friction coefficients on the diagonal.

In our example \( d = 2 \), we choose \( \beta = 1.67 \), \( \Gamma = I \) and we choose the triple well potential, given by

\[ V(x, y) = 3e^{-x^2-(y-\frac{1}{2})^2} - 3e^{-x^2-(y+\frac{1}{2})^2} - 5e^{-(x-1)^2-y^2} - 5e^{-(x+1)^2-y^2} + \frac{1}{5}x^4 + \frac{1}{5}(y - \frac{1}{3})^4, \tag{4.1} \]

on the state space \( S = \{(x, y) \in \mathbb{R}^2 \mid -2 \leq x \leq 2, -1.5 \leq y \leq 2.5\} \), as shown in Figure 1a. We employ reflecting boundary conditions.

To obtain the reactant and product state \( A \) and \( B \), we first define the basins of the energy landscape containing minima of the potential function \( \pm, 1, 0 \). In particular \( A' = \{(x, y) \in S \mid V(x, y) \leq -3, x \leq 0\} \) and \( B' = \{(x, y) \in S \mid V(x, y) \leq -3, x \geq 0\} \).

To discretise the state space \( S \) we use a particular Voronoi partition, namely a uniform \( 20 \times 20 \) mesh which yields squares with side length 0.2 as partition sets and a discrete state space \( I := \{1, 2, \ldots, 400\} \). The choice of this discretisation enables us to compare the results of our approach with the true values of the objects from TPT for diffusion processes, where these values are computed using the method of finite differences. The discrete reactant set \( A \) and product set \( B \) are approximations of the basins \( A' \) and \( B' \) respectively. The discretisation sets which form \( A \) and \( B \) are outlined in red in Figure 1b.

In order to simulate trajectories of the diffusion process \( X \) that solves the Smoluchowski equation with potential given by (4.1), we used the Euler-Maruyama method with time step \( \Delta t = 0.001 \). Sampled trajectories were then projected to the discrete state space \( I = \{1, 2, \ldots, 400\} \) in order to compute the discrete TPT objects of our approach, as well as the transition probability matrix of TPT for Markov chains.

To obtain the discrete committor function described in Section 3.1, we sampled \( 10^4 \) trajectories for each cell \( S_i \) in the partition. The initial condition of each trajectory was distributed uniformly in \( S_i \). Each sample trajectory was terminated when it first reached \( A \cup B \). To compute the discrete probability current \( \hat{J}_{AB,i} \) for each \( S_i \), \( i \in \{1, \ldots, 400\} \), we need to compute the vector \( \hat{\alpha}_i \in \mathbb{R}^{\#M'} \) of net flows of reactive trajectories, normalised by surface area; see (3.13) and the equations that precede it. Given the use of a uniform mesh into squares of the same length, the surface area measures of the form \( \sigma(\partial S_i \cap \partial S_k) \) for adjacent \( S_i \) and \( S_k \) were all equal, so computing \( \hat{\alpha}_i \) reduces to computing the \( \{\alpha_{ik}\}_{k \in \#M'} \). We did this by sampling a long trajectory of \( 5 \times 10^9 \) steps, and extracting from this trajectory approximately \( 10^8 \) segments that left \( A \) and reached \( B \) before returning to \( A \). These segments constitute our sampled reactive trajectories. The initial states of each reactive trajectory was distributed randomly on \( \partial A \).
(a) Potential function $V(x, y)$ given in (4.1) on the state space $S$.

(b) The potential $V(x, y)$ projected to the $(x, y)$ plane. The sets $A'$ and $B'$ are denoted in white. Partition sets that belong to the discrete sets $A$ and $B$ are outlined in red.

Figure 1: The energy landscape $V$.

To obtain the transition probability matrix of TPT for Markov chains, we sampled for each Voronoi cell $2 \times 10^5$ short trajectories. For each short trajectory, the initial condition was uniformly distributed on the cell. From this collection of short trajectories, we computed the transition probability matrix for a given lag time, for multiple choices of lag time. For each transition probability matrix, we computed the committor and the net flow of reactive trajectories, using the algorithm described in [16] and the PyEMMA package [21]. Note that what we refer to above as the ‘net flow of reactive trajectories’ corresponds to ‘probability current of reactive trajectories’ in [16]. We choose to use the former instead of the latter because in [16] the latter corresponds to a scalar-valued function on the facets of partition sets, instead of a vector-valued function on points in state space.

We computed the $L^2$ error of the committor function $\tilde{\eta}$ computed using our approach with respect to the true committor $\eta^{FD}$ computed using finite differences. For each lag time, we computed the $L^2$ error also of the committor $\eta^{MC}$ computed using TPT for Markov chains. We present our results in Figure 2a. Observe that as the lag time increases, the error for the committors of TPT for Markov chains decreases, until it reaches a minimum at about $\tau = 95\Delta t$. Any further increase in the lag time results in the slight increase in the error. Our approach does not depend on the lag time parameter, and is shown as a constant function for comparison.

Given the net flows $\{\alpha_{ik}\}_{k \in \mathbb{N}, i=1,\ldots,400}$, we computed the corresponding discrete probability currents using the normal equations (3.14); these were then used to compute the discrete streamlines of our approach according to (3.24) and (3.25). The streamlines for TPT for Markov chains were computed in the same way, with the difference that the net flows used in the normal equations were the normal flows computed for TPT for Markov chains using the PyEMMA method as described above. To compute the ‘ground truth’ streamlines, we first computed the probability current using finite differences. Since our diffusion process follows Smoluchowski dynamics, a closed-form expression for the probability current in terms of the gradient of the committor and the equilibrium density can be used [15]. We then obtained the streamlines by solving (3.22a) with the finite differences probability current in place of $J_{AB}$. In Figure 2b, we plot the streamlines computed using finite differences, using TPT for Markov chains, and using our approach, for 9 initial points on $\partial A$. 

18
Figure 2: Comparison of committors and streamlines generated by our approach and TPT for Markov chains with respect to ground truth value computed by finite differences.

It is known [18] that at the chosen temperature $\beta = 1.67$, the preferred transition channel is the ‘direct’ one that is orthogonal to the energetic barrier between the sets $A$ and $B$. The streamlines of both our approach and TPT for Markov chains approximate this channel equally well. On the other hand, for transition channels below or above the preferred one, we observe that the streamlines of TPT for Markov chains tend to deviate more from the ground truth than the streamlines of our approach.

5. Conclusion

In this paper, we applied ideas from TPT for diffusion processes taking values in a continuous, compact state space, in order to approximate certain objects by discrete analogues defined for a non-Markovian jump process on a finite state space. Our method uses the ideas that motivate the definitions of the committer, probability current, and streamline from TPT for diffusion processes, in order to construct analogues of these objects on the finite state space of the jump process. Our method transforms reactive trajectory data into objects that can be used to make predictions about the most likely paths that the diffusion process will take when transitioning from one metastable set to another. By giving error bounds and conditions under which the error bounds hold, we demonstrated the validity of our method, when the underlying process that generates the reactive trajectory data is an ergodic diffusion process. We presented numerical results that suggest that our method may be competitive with TPT for Markov chains in the computation of streamlines.

An important feature of our approach is that it does not require Markovianity, unlike TPT for Markov chains or for Markov jump processes. This feature is attractive, because it is often difficult to ensure in practice that an approximating process obtained from discretisation of a continuous state space possesses the Markov property. Furthermore, it is nontrivial to estimate errors due to making predictions or quantities that assume the Markov property when the
approximating object is non-Markovian. From the point of view of computational cost, not requiring Markovianity removes the need for identification of suitable lag times, which increases the amount of computational preprocessing necessary.

The approach we have taken here is statistical and data-driven in nature, in the sense that our method relies only on reactive trajectory data. This may be a disadvantage when reactive trajectory data is difficult to obtain. On the other hand, it is an advantage when the task of obtaining such data is easier, compared to the task of learning the drift and diffusion coefficients of the stochastic differential equation. Since it is based on sampling, our method suffers from the usual problems associated to sampling-based methods, e.g., the possibility of slow convergence of sample means to expected values. On the other hand, our method applies to problems in dimensions where standard, deterministic numerical methods for PDEs are too costly to be practical.

From the point of view of applications, there are two objects from transition path theory and similar theories that we have not considered here: the isocommittor surfaces, and reaction rates associated to a dividing surface between the reactant and product sets. Developing discrete analogues of these objects and describing how these discrete analogues may be approximated from reactive trajectory data would be useful for applications in molecular dynamics, for example.

The method and analysis that we have presented above can be extended in multiple directions. To simplify our error analysis, we omitted any discussion of the effect of statistical error in the predictions, and implicitly relied on the law of large numbers to ensure the convergence of statistics computed from finitely many samples to the true expected values. However, it would be of interest to estimate this statistical error, or to otherwise account for the uncertainty in the predictions made. It would be useful to relax certain assumptions, e.g., the compactness of the state space $S$, as well as the Lipschitz continuity of the committor and probability current. For applications where trajectory data is difficult to obtain, it would be of interest to extend the method so that it could be applied to sparse trajectory data, where ‘sparse’ means that subsequent measurement times are far apart from each other, or where the trajectory data degenerates to a cloud of point data. If the point data are assumed to lie on a manifold – in particular, on the graph of a potential – then such an extension might be possible. For example, a recent analysis [9] has shown that certain transition rate matrices constructed using point cloud data [11] converges in a particular sense to the infinitesimal generators of associated diffusion processes.

Acknowledgments

NC and TC are supported by project grant CH14 of the Einstein Center for Mathematics Berlin (ECMath), and by the DFG Research Center Matheon “Mathematics for key technologies” in Berlin. TC is supported by the German Ministry of Research and Education (BMBF) project grant 3FO18501 (Forschungscampus MODAL). HCL is supported by an open topic postdoctoral fellowship of Universität Potsdam.

References

[1] F. Aurenhammer. Voronoi diagrams: a survey of a fundamental geometric data structure. *ACM Comput. Surv.*, 23(3):345–405, September 1991.

[2] N. Boekelheide, R. Salomón-Ferrer, and T. F. Miller. Dynamics and dissipation in enzyme catalysis. *Proceedings of the National Academy of Sciences*, 2011.
[3] G. R. Bowman, V. A. Voelz, and V. S. Pande. Taming the complexity of protein folding. *Current Opinion in Structural Biology*, 21(1):4 – 11, 2011.

[4] 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, Aug 2014.

[5] N. Cvetković and H. C. Lie. On a degeneracy ratio for bounded convex polytopes, 2019. In preparation.

[6] Q. Du, V. Faber, and M. Gunzburger. Centroidal Voronoi Tessellations: Applications and Algorithms. *SIAM Review*, 41(4):637–676, 1999.

[7] W. E, W. Ren, and E. Vanden-Eijnden. Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes. *Chemical Physics Letters*, 413(1):242 – 247, 2005.

[8] W. E and E. Vanden-Eijnden. Towards a theory of transition paths. *Journal of statistical physics*, 123(3):503, 2006.

[9] Martin Heida. Convergences of the squareroot approximation scheme to the FokkerPlanck operator. *Mathematical Models and Methods in Applied Sciences*, 28(13):2599–2635, 2018.

[10] A. Klenke. *Probability theory: a comprehensive course*. Springer Science & Business Media, 2008.

[11] H. C. Lie, K. Fackeldey, and M. Weber. A square root approximation of transition rates for a Markov state model. *SIAM Journal on Matrix Analysis and Applications*, 34(2):738–756, 2013.

[12] J. Lu and J. Nolen. Reactive trajectories and the transition path process. *Probability Theory and Related Fields*, 161(1):195–244, Feb 2015.

[13] L. Maragliano and E. Vanden-Eijnden. On-the-fly string method for minimum free energy paths calculation. *Chemical Physics Letters*, 446(1):182 – 190, 2007.

[14] P. Metzner. *Transition path theory for Markov processes*. PhD thesis, Freie Universität Berlin, 2008.

[15] P. Metzner, C. Schütte, and E. Vanden-Eijnden. Illustration of transition path theory on a collection of simple examples. *The Journal of Chemical Physics*, 125(8):084110, 2006.

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

[17] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proceedings of the National Academy of Sciences*, 106(45):19011–19016, 2009.

[18] S. Park, M. K. Sener, D. Lu, and K. Schulten. Reaction paths based on mean first-passage times. *The Journal of Chemical Physics*, 119(3):1313–1319, 2003.
[19] N. Plattner and F. Noé. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nature Communications*, 6(7653), 2015.

[20] J. Prinz, B. Keller, and F. Noé. Probing molecular kinetics with Markov models: metastable states, transition pathways and spectroscopic observables. *Phys. Chem. Chem. Phys.*, 13, 2011.

[21] M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J. Prinz, and F. Noé. PyEMMA 2: A software package for estimation, validation, and analysis of Markov models. *Journal of Chemical Theory and Computation*, 11:5525–5542, October 2015.

[22] R. Schneider and J. A. Wieacker. Approximation of convex bodies by polytopes. *The Bulletin of the London Mathematical Society*, 13(2):149–156, 1981.

[23] D. Silva, G. R. Bowman, A. Sosa-Peinado, and X. Huang. A role for both conformational selection and induced fit in ligand binding by the lao protein. *PLOS Computational Biology*, 7(5):1–11, 05 2011.

[24] A. Sznitman. *Brownian motion, obstacles and random media*. Springer Science & Business Media, 1998.

[25] B. Trendelkamp-Schroer and F. Noé. Efficient estimation of rare-event kinetics. *Phys. Rev. X*, 6:011009, Jan 2016.

[26] E. Vanden-Eijnden. *Transition Path Theory*, pages 91–100. Springer Netherlands, Dordrecht, 2014.

[27] E. Vanden-Eijnden and M. Venturoli. Markovian milestoning with Voronoi tessellations. *The Journal of Chemical Physics*, 130(19):194101, 2009.

[28] E. Vanden-Eijnden, M. Venturoli, G. Ciccotti, and R. Elber. On the assumptions underlying milestoning. *The Journal of Chemical Physics*, 129(17):174102, 2008.

[29] M. von Kleist, C. Schütte, and W. Zhang. Statistical analysis of the first passage path ensemble of jump processes. *Journal of Statistical Physics*, 170(4):809–843, 2018.

### A. Basic facts about polytopes

**Proposition A.1.** Let $d \in \mathbb{N}$, $d > 1$. Then any $d$-dimensional polytope in $\mathbb{R}^d$ has at least $d + 1$ facets.

**Proof.** We prove the claim by induction.

**Base case:** Let $d = 2$. The polytope in $\mathbb{R}^2$ of full dimension with the smallest number of facets is a triangle, which has $3 = d + 1$ facets.

Suppose that the claim holds for a $(d-1)$-dimensional polytope.

**Induction step:** Let $d \geq 3$, and assume that there exists an $d$-dimensional polytope with only $d$ facets. These $d$ facets are $(d-1)$-dimensional polytopes. Furthermore, each facet of the original $d$-dimensional polytope intersects at most $d - 1$ other facets, since there are $d$ facets in total by assumption. This yields that there exist $(d-1)$-dimensional polytopes with at most $d - 1$ facets, which contradicts the base case. Thus any $d$-dimensional polytope must have at least $d + 1$ facets. \(\square\)
Corollary A.2. Let $P$ be a $d$-dimensional polytope in $\mathbb{R}^d$. Then $P$ has at least $d$ linearly independent outer normals.

Proof. By Proposition A.1, $P$ has at least $d+1$ facets, and therefore at least $d+1$ outer normals. We will prove by contradiction that there exist $d$ linearly independent outer normals.

Suppose $P$ has no more than $d - 1$ linearly independent outer normals. Then the normals to the facets of $P$ span at most a $(d-1)$-dimensional space, which implies that there exists a hyperplane $H$ in $\mathbb{R}^d$ containing all the outer normals of $P$. Let $v$ be normal to $H$, and let $n$ be an arbitrary outer normal associated to some facet $F$ of $P$. Then $v$ and $n$ are orthogonal, which implies that $v$ is parallel to $F$, and thus that $F$ is unbounded along the direction of $v$. This implies that $P$ is unbounded, which produces the desired contradiction. \qed