Convergent discretisation schemes for Transition Path Theory for diffusion process

Nada Cvetković[*], Tim Conrad, and Han Cheng Lie

Department of Mathematics and Computer Science, Freie Universität Berlin, Arnimallee 6, 14195 Berlin, 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, the applicability of TPT to diffusion processes is often limited by the fact that many methods for computing the primary objects from TPT involve solving a partial differential equation (PDE) by mesh-based methods, and thus suffer from the curse of dimensionality. We describe novel mesh-free discretisation schemes for committors and isocomittor surfaces, as well as probability currents and their streamlines, i.e. transition paths. Our method uses a Voronoi tesselation and simple Monte Carlo to circumvent the curse of dimensionality; this leads to an approximation of the underlying diffusion process by a non-Markovian jump process on the Delaunay graph dual to the Voronoi tesselation. We prove rigorous error bounds and convergence theorems for the discretised objects, thus ensuring the validity of our approach.

1. Introduction

Transition path theory (TPT) is a statistical theory developed for the analysis of transition paths of a Markov process between two subsets of state space [6, 7]. TPT was constructed for processes for the subsets of interest are metastable, i.e. when transitions between the subsets are rare. 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 with [11, 14, 22, 3, 17, 2, 15, 16, 19, 20] providing a far from exhaustive list of references in this area. TPT has also been used to study flows in complex networks [4]. TPT has also been developed for Markov jump processes [12] and studied using the tools of stochastic analysis [10].

One of the most fundamental objects in TPT is the forward committor functions, which are also known as capacitors in probability theory [18]. Let S denote the state space of the Markov process, and let A and B denote the two subsets of interest, whereby we are interested in transitions from A to B. Then the *forward committor function* $q : S \setminus (A \cup B) \to [0, 1]$ is a

[*]Corresponding author: ncvetkovic@math.fu-berlin.de

function such that for any $x \in S \setminus (A \cup B)$, $q(x)$ gives the probability of reaching A before B, given that the current state of the Markov process is x,

$$q(x) = \mathbb{P}(X_{\tau_{A\cup B}} \in B | X_0 = x), \tag{1.1}$$

where $\tau_{A\cup B}$ denotes the *first hitting time* of the process $(X_t)_{t>0}$ with respect to set $A \cup B$, i.e. $\tau_{A\cup B} = \inf\{t \geq 0 | X_t \in A \cup B\}$.

One of the most powerful aspects of TPT is that it 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. The power of these relationships is 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, and apply deterministic methods to compute all the objects from TPT with which one can describe transitions between A and B. This can be done efficiently, and can lead to useful visualizations of the behaviour of reactive trajectories, see e.g. [13].

A key weakness of TPT is that the powerful relationships it describes are often difficult to exploit in practice. This is because complex systems are often high-dimensional, and the curse of dimensionality renders virtually all deterministic methods for solving partial differential equations too costly to be practical. This creates a need for methods that can apply TPT to study complex systems, but that are not based on solving partial differential equations.

The goal of this paper is to describe a way to discretise the objects from TPT in such a way that the discretisations do not suffer from the curse of dimensionality, for diffusion processes taking values in bounded subsets of $\mathbb{R}^d$. We do so by using Voronoi tesselations, which are a powerful discretisation scheme that have been applied in a wide range of contexts [5, 1]. Voronoi tesselations have also been used to circumvent the curse of dimensionality in the context of molecular dynamics [21, 9]. Our discretisation method leads to a non-Markovian jump process approximation of the underlying continuous state space Markov process. We describe how the Voronoi tesselation leads to this non-Markovian jump process model in Section 2. In Section 3 we define discrete analogues of the committor and isocommittor surfaces, as well as probability currents and the associated streamlines, for this non-Markovian jump process. We show that these discrete analogues provide increasingly accurate approximations of the corresponding objects for the underlying diffusion process, as the discretisations provided by the Voronoi tesselation become finer.

2. Setup

Let $X = (X_t)_{t\geq 0}$ be an ergodic diffusion process taking values in a bounded subset $S \subset \mathbb{R}^d$. Suppose that the invariant measure μ of X is absolutely continuous with respect to Lebesgue measure, so that there exists a density $p : S \to \mathbb{R}$ such that

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

A Voronoi tesselation 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

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

That is, S_i is the closed set consisting of all points in state space that are closer 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. Observe that

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

so that Voronoi cells intersect at most at their boundaries; note that this intersection may be empty. This motivates 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 tesselation $\{S_i\}_{i\in I}$, 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

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

This leads us to the next definition.

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

$$\rho(\{S_i\}_{i\in I}) := \max_{i\in I} \operatorname{diam}(S_i).$$

When there is no risk of confusion, we will omit the argument $\{S_i\}_{i\in I}$ of the width and simply write ρ. The smaller (resp. larger) the width, the finer (resp. coarser) the tesselation. We shall be interested in obtaining error bounds in the limit of small width, i.e. as ρ decreases to 0. As ρ decreases to zero, the number n of cells in the Voronoi tesselation must increase to infinity. The converse is not true: it is possible for the number of cells to increase to infinity, while the width ρ stays bounded away from zero. The main idea that we will exploit is that as ρ decreases to zero, all the Voronoi cells shrink to their associated generators.

Given a Voronoi tesselation $\{S_i\}_{i\in I}$ with finite index set I and its associated Delaunay graph $G = (I, E)$, we construct a non-Markovian jump process $Y = (Y_t)_{t\geq 0}$ on the index set I, by setting $Y_t = i$ whenever $X_t \in S_i$. For the case that $X_t \in S_i \cap S_j$ for a pair (S_i, S_j) of adjacent Voronoi cells, then we set $Y_t = k \in \{i, j\}$, where k is the index set of the cell such that the diffusion process X was 'most recently' in the interior of S_k prior to time t. To be precise, if there exists some $\epsilon > 0$ such that $X_s \in S_k$ for all $t - \epsilon < s < t$, then we set $Y_t = k$.

In this paper, we write $A, B \subset S$ to denote the reactant and product sets of states for the diffusion process of interest. We will assume that A and B are simply connected, compact and disjoint.

3. Convergence analysis of TPT

In this section we define transition path theory objects for continuous processes on a discrete state space and we show that in the limit of small partition size the results of the discrete path theory converge to the results of transition path theory for diffusion processes. More precisely, we define the discrete objects such as discrete committors, discrete probability current and discrete streamlines and prove their convergence, with respect to an appropriate metric, to the respective continuous objects of diffusion processes.

3.1. Committors

Recall the forward committor function defined in (1.1). For $i \in I$, define

$$\hat{q}_i = \frac{1}{\mu(S_i)} \langle q, \mathbf{1}_{S_i} \rangle_\mu, \tag{3.1}$$

where $\mathbf{1}_{S_i} : S \to \{0, 1\}$ is the indicator function of S_i and $\langle \cdot, \cdot \rangle_\mu$ is an inner product with respect to the invariant measure μ of the process X_t, i.e. $\langle v, w \rangle_\mu = \int_S v(x)w(x)\mu(\mathrm{d}x)$.

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 forward committor function* $\hat{q} : S \to [0, 1]$:

$$\hat{q}(x) := \sum_{i \in I} \hat{q}_i \mathbf{1}_{\mathrm{int}(S_i)}(x). \tag{3.2}$$

To complete the definition of $\hat{q}$, we need to specify its values on the intersections of the closed 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 to q in an L^p norm, the values that we prescribe will not be important. One straightforward assignment is as follows: If $x \notin \mathrm{int}(S_i)$ for all $i \in I$, then there must be a nonempty set $C \subset I$ such that $x \in \partial S_c$ for all $c \in C$; 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 mean of $\hat{q}_c$ over $c \in C$.

Now we define the *discrete forward committor function* $\tilde{q} : \{1, \ldots, n\} \to [0, 1]$, which is the forward committor function that corresponds to the continuous process $(Y_t)_{t \geq 0}$ on the discrete state space. We will make the following assumption.

Assumption 3.1. *There exist disjoint subsets $J, K \subset \{1, \ldots, n\}$ such that $A = \cup_{j \in J} S_j$ and $B = \cup_{k \in K} S_k$.*

Thus, J and K are the metastable sets for the jump process Y that correspond to the metastable sets A and B for the diffusion process X. The assumption is justified in the limit of small width, since one can approximate any simply connected, compact subset of state space arbitrarily well by collections of arbitrarily small Voronoi cells.

Given J and K, define the first hitting time of the jump process to $J \cup K$

$$\tau_{J \cup K} := \inf\{t \geq 0 \mid Y_t \in J \cup K\}. \tag{3.3}$$

Given our definition of Y, J, and K, it holds that the first hitting time $\tau_{J \cup K}$ of the jump process Y and the first hitting time $\tau_{A \cup B}$ of the diffusion X coincide.

We define the discrete forward committor by the conditional probability

$$\tilde{q}_i = \frac{\mathbb{P}\left(Y_{\tau_{J \cup K}} \in K,\ Y_0 = i\right)}{\mathbb{P}\left(Y_0 = i\right)}. \tag{3.4}$$

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 3.2 (Stochastic kernel)**.** Let $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ be measurable spaces. A map $\kappa : \Omega_1 \times \mathcal{A}_2 \to [0, \infty]$ is called a *stochastic kernel* from Ω_1 to Ω_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 $\omega_1 \in \Omega_1$.

Definition 3.3 (Regular conditional distribution)**.** Let $(\Omega_1, \mathcal{A}_1, \mathbb{P})$ be a probability space, $(E, \mathcal{E})$ be a measurable space, and Y be an E-valued random variable on $(\Omega_1, \mathcal{A}_1, \mathbb{P})$. Let $\mathcal{G} \subset \mathcal{A}_1$ be a σ-algebra on Ω_1. A stochastic kernel $\kappa_{Y,\mathcal{G}}$ from $(\Omega_1, \mathcal{G})$ to $(E, \mathcal{E})$ is called a *regular conditional distribution of* Y *given* $\mathcal{G}$ if

$$\kappa_{Y,\mathcal{G}}(\omega, C) = \mathbb{P}\left(Y \in C|\mathcal{G}\right)(\omega) \tag{3.5}$$

for $\mathbb{P}-$almost all $\omega \in \Omega_1$ and for all $C \in \mathcal{E}$.

If $\mathcal{G}$ is generated by a random variable X defined on $(\Omega_1, \mathcal{A}_1, \mathbb{P})$ that takes values in some measurable space $(E', \mathcal{E}')$ then the stochastic kernel $\kappa_{Y,X}$ from $(E', \mathcal{E}')$ to $(E, \mathcal{E})$ defined by

$$\kappa_{Y,X}(x, C) = \mathbb{P}\left(Y \in C|X = x\right) = \kappa_{Y,\sigma(X)}(X^{-1}(x), C) \tag{3.6}$$

is called a *regular conditional distribution of* Y *given* X.

The existence of the regular conditional distribution of Y given $\mathcal{G}$ is shown in [8, p. 185]. The existence of the regular conditional distribution of Y given X requires the factorisation lemma [8, p. 38].

Remark 3.4 (Committors are regular conditional probabilities)**.** *By setting* $Y = X_{\tau_{A\cup B}(X)}$*,* $X = X_0$*, and* $C = B$ *in* (3.6)*, we obtain the definition* (1.1) *of the forward committor probability.*

The following theorem is given in [8, p. 185]

Theorem 3.5 (Conditional expectations in terms of regular conditional distributions)**.** *Let* Y *be a random variable on* $(\Omega_1, \mathcal{A}_1, \mathbb{P})$ *with values in some set* E*, and equip* E *with the Borel* σ*-algebra* $\mathcal{E}$*. Let* $\mathcal{G} \subset \mathcal{A}_1$ *be a* σ*-algebra and let* $\kappa_{Y,\mathcal{G}}$ *be a regular conditional distribution of* Y *given* $\mathcal{G}$*. Further, let* $f : E \to \mathbb{R}$ *be measurable and* $\mathbb{E}\left[|f(Y)|\right] < \infty$*. Then*

$$\mathbb{E}\left[f(Y)|\mathcal{G}\right](\omega) = \int_E f(y)\kappa_{Y,\mathcal{G}}(\omega, \mathrm{d}y)$$

$\mathbb{P}$*-almost surely.*

We have the following corollary.

Corollary 3.6. *Let* X *and* Y *be random variables on* $(\Omega_1, \mathcal{A}_1, \mathbb{P})$*, taking values in the measurable spaces* $(E', \mathcal{E}')$ *and* $(E, \mathcal{E})$ *respectively. Then for any* $C \in \mathcal{E}$ *and* $D \in \mathcal{E}'$*,*

$$\mathbb{P}\left(Y \in C, X \in D\right) = \int_D \kappa_{Y,X}(x, C)\mathbb{P} \circ X^{-1}(\mathrm{d}x).$$

Proof. Fix an arbitrary $\omega \in \Omega_1$. Since probabilities of events are expected values of indicator functions, the properties of conditional expectations yield

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

Consider the conditional probability inside the expectation of the last term. Setting $\mathcal{G} = \sigma(X)$ and $f = \mathbf{1}_C$ in Theorem 3.5, and using (3.6), we obtain

$$\mathbb{P}(Y \in C|X)(\omega) = \mathbb{E}[\mathbf{1}_\mathbf{C}(Y)|X](\omega) = \int_E \mathbf{1}_C(y)\kappa_{Y,\sigma(X)}(\omega, \mathrm{d}y) = \kappa_{Y,\sigma(X)}(\omega, C) = \kappa_{Y,X}(X(\omega), C).$$

Therefore,

$$\begin{aligned}\mathbb{E}[\mathbb{P}(Y \in C|X)\mathbf{1}_D(X)] &= \int_{\Omega_1} \mathbf{1}_D(X(\omega))\kappa_{Y,X}(X(\omega),C)\mathbb{P}(\mathrm{d}\omega)\\ &= \int_{E'} \mathbf{1}_D(x)\kappa_{Y,X}(x,C)\mathbb{P}\circ X^{-1}(\mathrm{d}x)\end{aligned}$$

as desired. □

We now use Corollary 3.6.

Proposition 3.7. *Suppose that Assumption 3.1 holds. Let the projected forward committor $\hat{q}_i$ and discrete forward committor $\tilde{q}_i$ be defined as in (3.1) and (3.4), respectively. Assume that X_0 is distributed according to the equilibrium distribution μ. Then $\hat{q}_i = \tilde{q}_i$, for all $i \in \{1, \dots, n\}$.*

Proof. Let $i \in \{1, \dots, n\}$ be arbitrary. Recall from (3.1) that

$$\hat{q}_i = \frac{1}{\mu(S_i)}\langle q, \mathbf{1}_{S_i}\rangle_\mu = \frac{1}{\mu(S_i)}\int_S q(x)\mathbf{1}_{S_i}(x)\mu(\mathrm{d}x).$$

The definition (3.4), the construction of Y and of the sets J and K, the definition (3.3), and the hypothesis that X_0 is distributed according to the equilibrium measure μ imply that

$$\tilde{q}_i = \frac{\mathbb{P}\left(Y_{\tau_{J\cup K}} \in K,\ Y_0 = i\right)}{\mathbb{P}\left(Y_0 = i\right)} = \frac{\mathbb{P}\left(X_{\tau_{A\cup B}} \in B,\ X_0 \in S_i\right)}{\mu(S_i)} = \frac{\mathbb{P}\left(X_{\tau_{A\cup B}} \in B,\ X_0 \in S_i\right)}{\mu(S_i)}.$$

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

$$\int_S q(x)\mathbf{1}_{S_i}(x)\mu(\mathrm{d}x) = \mathbb{P}\left(X_{\tau_{A\cup B}} \in B,\ X_0 \in S_i\right).$$

By Remark 3.4, the left-hand side can be rewritten in terms of a regular conditional probability,

$$\int_S q(x)\mathbf{1}_{S_i}(x)\mu(\mathrm{d}x) = \int_S \mathbf{1}_{S_i}(x)\kappa_{X_{\tau_{A\cup B}(X)},X_0}(x,B)\mu(\mathrm{d}x) = \int_{S_i} \kappa_{X_{\tau_{A\cup B}(X)},X_0}(x,B)\mu(\mathrm{d}x).$$

Using that $\mu = \mathbb{P}\circ X_0^{-1}$ and Corollary 3.6, we obtain

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

yielding the desired conclusion. □

The following lemma uses the convexity of Voronoi cells and the continuity of the forward committor function. We shall use the lemma later to prove Theorem 3.9.

Lemma 3.8. *Let $\{S_i\}_{i\in I}$ be a Voronoi tesselation 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 $\hat{q}_i$, and so there exist uncountably many x_i which satisfy the desired property. Therefore, suppose that q is not constant on S_i, and partition S_i into the disjoint subsets $S_i^- := \{x \in S_i \ :\ q(x) < \hat{q}|_{S_i}\}$, $S_i^+ := \{x \in S_i \ :\ q(x) > \hat{q}|_{S_i}\}$ and $S_i^0 := \{x \in S_i \ :\ q(x) = \hat{q}|_{S_i}\}$. 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. □

Next, we show that our choice of the projected committor function (3.1) is valid, by proving an error bound for the error incurred when approximating the true forward committor q with the projected forward committor $\hat{q}$ defined in (3.2).

Theorem 3.9. *Suppose that the forward committor $q : S \to [0,1]$ has bounded derivatives of first order, i.e. $\nabla q \in L^\infty$, and let $p \in [1,\infty)$. Then there exists some $C > 0$ that depends only on q, such that for any Voronoi tesselation $\{S_i\}_{i\in I}$ of S with width ρ, the corresponding projected committor function $\hat{q}$ satisfies*

$$\|q - \hat{q}\|_p \le C\rho$$

In particular, as the width of the Voronoi tesselation decreases to zero, the L^p error of $\hat{q}$ decreases linearly with ρ.

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 $i \in I$. By Lemma 3.8, there exists an $x_i \in S_i$ such that $q(x_i) = \hat{q}_i = \hat{q}|_{S_i}$.

Now, consider the first order Taylor expansion of q about the point x_i:

$$q(x) = q(x_i) + \langle \nabla q(x_i), x - x_i \rangle + \mathcal{O}\left(|x - x_i|\right), \quad x \in S_i.$$

Computing the L^p-error of the restrictions of q and $\hat{q}$ to S_i, using that $\hat{q}|_{S_i} = q(x_i)$ by definition of x_i, and using the Taylor expansion above, we obtain

$$\begin{aligned} \|\left(q - \hat{q}\right)|_{S_i}\|_p^p &= \int_{S_i} (q(x) - \hat{q}(x))^p \,\mathrm{d}x = \int_{S_i} (q(x) - q(x_i))^p \,\mathrm{d}x \\ &= \int_{S_i} \left(\langle \nabla q(x_i), x - x_i\rangle + \mathcal{O}\left(|x - x_i|\right)\right)^p \mathrm{d}x \\ &\le 2^{p-1} \int_{S_i} \left(|\nabla q(x_i)|\,|x - x_i|\right)^p + \mathcal{O}\left(|x - x_i|^p\right) \mathrm{d}x, \end{aligned}$$

where we used the Cauchy-Schwarz inequality and the inequality $(a+b)^p \le 2^{p-1}(a^p + b^p)$ for the last inequality. Since $x, x_i \in S_i$ implies that $|x - x_i| \le \rho$, we have

$$\int_{Si} (|\nabla q(x_i)|\,|x - x_i|)^p \mathrm{d}x \le (\|\nabla q\|_\infty \rho)^p \mu(S_i).$$

Therefore, to leading order, we have that

$$\|q - \hat{q}\|_p^p = \sum_{i\in I} \|(q - \hat{q})|_{S_i}\|_p^p \le 2^{p-1} \|\nabla q\|_\infty^p \rho^p \left(\sum_{i\in I} \mu(S_i)\right) \le (2\|\nabla q\|_\infty \rho)^p,$$

where the last inequality follows from $2^{p-1} \le 2^p$ and $\sum_{i\in I} \mu(S_i) = 1$. This completes the proof, with $C = 2\|\nabla q\|_\infty$. ☐

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

Corollary 3.10. *Suppose that Assumption 3.1 and the assumptions of Theorem 3.9 holds. Then for the same scalar C, it holds that for any Voronoi tesselation $\{S_i\}_{i\in I}$ of S with width ρ, the function $\tilde{q}$ satisfies*

$$\|q - \tilde{q}\|_p \le C\rho,$$

and the L^p error of $\tilde{q}$ decreases linearly with the width ρ.

Proof. The result follows from Theorem 3.9 and Proposition 3.7. ☐

3.2. Isocommittor surfaces

For a given value a in the unit interval $[0, 1]$, the corresponding isocommittor surface is the a-level set of the true forward committor q, i.e.

$$q_a := \{x \in S \ : \ q(x) = a\}.$$

Then by definition of q, $q_0 = A$, and $q_1 = B$.

Assumption 3.11. *The true forward committor q defined in (1.1) is globally Lipschitz continuous with constant $K > 0$, i.e.*

$$|q(x) - q(y)| \leq K |x - y| , \quad \forall x, y \in S.$$

Fix a Voronoi tesselation $\{S_i\}_{i \in I}$ with width ρ. Recall the definitions (3.4) of $\tilde{q}_i$ and $\hat{q}_i$ respectively, and that $\tilde{q}_i = \hat{q}_i$ by Proposition 3.7. We shall use the following lemma.

Lemma 3.12. *Suppose that Assumption 3.11 holds. Let $\{S_i\}_{i \in I}$ be a Voronoi tesselation with width ρ. Then for any $i \in I$, it holds that*

$$\tilde{q}_i - K\rho \leq q(y) \leq \tilde{q}_i + K\rho, \quad \forall y \in S_i.$$

Proof. Fix an arbitrary $i \in I$. Given Assumption 3.1, it holds that

$$|q(x) - q(y)| \leq K |x - y| \leq K\rho, \quad \forall x, y \in S_i. \tag{3.7}$$

Now recall that, by Lemma 3.8, there exists at least one $x_i \in S_i$ such that $q(x_i) = \hat{q}_i$. Hence by Proposition 3.7, $q(x_i) = \tilde{q}_i$. Therefore, we have

$$|\tilde{q}_i - q(y)| = |q(x_i) - q(y)| \leq K |x_i - y| \leq K\rho,$$

which proves the claim. □

Suppose we have a Voronoi tesselation $\{S_i\}_{i \in I}$ of width ρ, together with the associated set $\{\tilde{q}_i\}_{i \in I}$ of discrete committor values. For a given a in the unit open interval, we define the corresponding *discrete isocommittor surface* $\tilde{q}_a(\rho)$ by the index set I_a and corresponding subset $\tilde{q}_a$ of state space,

$$I_a := \{i \in I \ : |\tilde{q}_i - a| \leq K\rho\}, \quad \tilde{q}_a(\rho) := \cup_{i \in I_a} S_i. \tag{3.8}$$

This definition is motivated by Corollary 3.10, which suggests that as the width decreases, there should be indices $i \in I$ such that the deviation of $\tilde{q}_i$ from a should not be too large.

The following lemma suggests that our definition is reasonable.

Lemma 3.13. *Fix an arbitrary a in the open unit interval, and let q_a denote the corresponding continuous isocommittor surface. Then $q_a \subset \tilde{q}_a(\rho)$.*

Proof. Let S_i be a partition set such that $S_i \cap q_a \neq \emptyset$, and suppose that $i \notin I_a$. According to definition (3.8), if $i \notin I_a$, then either $\tilde{q}_i < a - K\rho$ or $\tilde{q}_i > a + K\rho$. If $\tilde{q}_i < a - K\rho$, then the right inequality of Lemma 3.12 yields that $q(y) \leq \tilde{q}_i + K\rho < a$, for all $y \in S_i$. This implies that $q_a \cap S_i = \emptyset$, which contradicts the assumption that $S_i \cap q_a$ is nonempty. Similarly, if $\tilde{q}_i > a + K\rho$, then the left inequality of Lemma 3.12 yields that $a < \tilde{q}_i - K\rho \leq q(y)$ for all $y \in S_i$, which again produces a contradiction with the assumption that $S_i \cap q_a$ is nonempty. Thus, if $S_i \cap q_a$ is nonempty, then i must belong to I_a. By definition of the discrete isocommittor $\tilde{q}_a(\rho)$, it follows that $S_i \subset \tilde{q}_a(\rho)$.

Since S_i was taken to be an arbitrary partition set that had nonempty intersection with q_a, it follows that every such partition set is contained in $\tilde{q}_a(\rho)$. Since q_a is contained in the union of sets S_i such that $S_i \cap q_a$ is nonempty, and since every such S_i is contained in $\tilde{q}_a(\rho)$, it follows that q_a is contained in $\tilde{q}_a(\rho)$. □

We now state our convergence result for the discrete isocommittor.

Theorem 3.14. *Suppose that Assumption 3.11 holds. Let $a \in [0,1]$ be arbitrary, and let $(\rho_n)_{n\in\mathbb{N}} \subset (0,+\infty)$ be a sequence that decreases to zero. For each ρ_n, consider an arbitrary Voronoi tesselation of S with width ρ_n, and let $(\tilde{q}_a(\rho_n))_{n\in\mathbb{N}}$ be the associated sequence of discrete isocommittor surfaces. Then there exists a family of sets $(D_a(\rho_n))_{n\in\mathbb{N}}$ such that $D_a(\rho_n) \supset \tilde{q}_a(\rho_n)$ and $(D_a(\rho_n))_{n\in\mathbb{N}}$ decreases to q_a.*

Proof of Theorem 3.14. It suffices to prove the first statement, since the second follows from the first statement and Lemma 3.13. Let $a \in [0,1]$ be arbitrary. Fix an arbitrary Voronoi tesselation $\{S_i\}_{i\in I}$ with width ρ, let I_a be as in (3.8), and fix an arbitrary $i \in I_a$. Then it holds that

$$a - K\rho \leq \tilde{q}_i \leq a + K\rho.$$

Combining these inequalities with the result of Lemma 3.12 yields

$$a - 2K\rho \leq q(y) \leq a + 2K\rho, \quad \forall y \in S_i.$$

Therefore, it holds that

$$|q(y) - a| \leq 2K\rho, \quad \forall y \in S_i,\ \forall i \in I_a. \tag{3.9}$$

For some $\rho' > 0$ not necessarily equal to ρ, define the set

$$D_a(\rho') := \{x \in S \ : \ |q(x) - a| \leq 2K\rho'\}.$$

From the definition of q_a and $D_a(\rho')$, it follows that $q_a \subset D_a(\rho')$, for any $\rho' > 0$. From (3.9), it follows that $S_i \subset D_a(\rho)$ for all $i \in I_a$. Thus

$$\tilde{q}_a(\rho) := \bigcup_{i\in I_a} S_i \subset D_a(\rho).$$

By Lemma 3.13 , we therefore have

$$q_a \subset \tilde{q}_a(\rho) \subset D_a(\rho). \tag{20}$$

To complete the proof, let $(\rho_n)_{n\in\mathbb{N}}$ be an arbitrary decreasing sequence of strictly positive numbers. For each $n \in \mathbb{N}$, consider a Voronoi tesselation with width ρ_n, the associated discrete isocommittor surface $\tilde{q}_a(\rho_n)$, and the associated set $D_a(\rho_n)$. It follows from the definition of $D_a(\rho')$ above that $D_a(\rho_{n+1}) \subset D_a(\rho_n)$, so $(D_a(\rho_n))_{n\in\mathbb{N}}$ decreases to q_a, and by (20) it follows that $\tilde{q}_a(\rho_n) \subset D_a(\rho_n)$ for all $n \in \mathbb{N}$, as desired. □

3.3. Probability current

In this section we define a *discrete probability current* $\tilde{J}_{AB}$ which is obtained by observing the diffusion process in the discretized 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 continuous process $(Y_t)_{t\geq 0}$. This means that, for any $t \geq 0$, we can detect which Voronoi cell S_i contains X_t. However, we cannot detect the exact location of X_t.

The probability current $J_{AB} : S \to \mathbb{R}^d$ represents the probability of the flow of reactive trajectories. On any surface ∂S_i, which is the boundary of $S_i \subset S \setminus (A \cup B)$, J_{AB} is defined implicitly via

$$\begin{aligned}\lim_{s\to 0^+} \frac{1}{s} \lim_{T\to\infty} \frac{1}{2T} \int_{\mathcal{R}\cap[-T,T]} \mathbf{1}_{S_i}\left(X(t)\right) \mathbf{1}_{S_i^{\complement}}\left(X(t+s)\right) - \mathbf{1}_{S_i^{\complement}}\left(X(t)\right) \mathbf{1}_{S_i}\left(X(t+s)\right) \mathrm{d}t \\ = \int_{\partial S_i} J_{AB}(y) \cdot n_{S_i}(y) d\sigma(y)\end{aligned}$$

where $n_{S_i}(y)$ denotes the outer unit normal to S_i at some point $y \in \partial S_i$. Notice that for a Voronoi cell S_i we have

$$\int_{\partial S_i} J_{AB}(y) \cdot n_{S_i}(y) d\sigma(y) = \sum_{k\in\mathcal{N}_i} \int_{\partial S_i \cap \partial S_k} J_{AB}(y) \cdot n_{ik}(y) \mathrm{d}\sigma(y), \tag{3.10}$$

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

$$\begin{aligned}\alpha_{i,j} &:= \lim_{s\to 0^+} \frac{1}{s} \lim_{T\to\infty} \frac{1}{2T} \int_{\mathcal{R}\cap[-T,T]} \mathbf{1}_{S_i}\left(X(t)\right) \mathbf{1}_{S_k}\left(X(t+s)\right) - \mathbf{1}_{S_k}\left(X(t)\right) \mathbf{1}_{S_i}\left(X(t+s)\right) \mathrm{d}t \\ &= \int_{\partial S_i \cap \partial S_k} J_{AB}(y) \cdot n_{ik} \mathrm{d}\sigma(y).\end{aligned} \tag{3.11}$$

To define the discrete probability current $\tilde{J}_{AB} : S \to \mathbb{R}^d$, we shall adopt the same strategy as we did for discrete forward committor $\tilde{q}$ defined in (3.4): we shall assume that the discrete probability current is piecewise constant and equal to some vector $\tilde{J}_{AB,i} \in \mathbb{R}^d$ on the interiors of each Voronoi cell $\mathrm{int}(S_i)$, and then define the value of $\tilde{J}_{AB}$ on facets. Choosing a piecewise constant function motivates the relation

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

Since we can observe the process $(Y_t)_{t\geq 0}$, we can approximate the quantity $\alpha_{i,j}$ by using sample data from reactive trajectories. Combining (3.11) with the preceding equation, it follows that for every $k \in \mathcal{N}_i$ we want to have

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

In the matrix notation, we can rewrite the above as a matrix-value equation

$$M_i \tilde{J}_{AB,i} = \hat{\alpha}_i \in \mathbb{R}^{\#\mathcal{N}_i}, \tag{3.12}$$

where $\hat{\alpha}_i$ is a vector of length $\#\mathcal{N}_i$ with k-th entry given by $(\hat{\alpha}_i)_k = \alpha_{i,k}/\sigma(\partial S_i \cap \partial S_k)$, and M_i is a $(\#\mathcal{N}_i) \times d$ matrix with real entries such that the k-th row is n_{ik}^T. A necessary condition for (3.12) to define a vector $\tilde{J}_{AB,i}$ is that $\hat{\alpha}_i$ must belong to the column space of N_i. However, we do not expect the latter condition to hold in general, since Proposition A.1 indicates that $\#\mathcal{N}_i \geq d+1$, 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 3.15 (N_i has full rank)**.** *Let $\{S_i\}_{i\in I}$ be a Voronoi tesselation of S, let $i \in I$ be arbitrary, and let n_{ik} be the unit normal on $\partial S_i \cap \partial S_k$ exterior to S_i. Then the corresponding matrix $N_i \in \mathbb{R}^{(\#\mathcal{N}_i)\times d}$ 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 $\#\mathcal{N}_i$ is at least $d+1$, i.e. S_i has at least $d+1$ adjacent facets, and thus 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. □

The lemma above implies that we can choose d rows from M_i in order to form an invertible matrix $N_i \in \mathbb{R}^{d\times d}$. Using the notation A_i to denote the i-th column of a matrix A and $\sigma_{\min}(A)$ (resp. $\sigma_{\max}(A)$) to denote the smallest (resp. largest) singular value of A, we define

$$N_i = \arg\max\{\sigma_{\min}(N_i') \ : \ \tilde{N}_i \in \mathbb{R}^{d\times d},\ \forall k \in [d],\ \exists j \in \mathcal{N}_i \text{ such that } (\tilde{N}_i^\top)_k = (M_i^\top)_j\}, \quad (3.13)$$

i.e. N_i is any square matrix that maximises the smallest singular value among all all square matrices with rows taken from the rows of M_i. Since N_i is a function of S_i, we may consider the smallest and largest singular values $\sigma_{\min}(N_i)$ and $\sigma_{\max}(N_i)$ of N_i, i.e.

$$\sigma_{\min}(S_i) := \sigma_{\min}(N_i), \quad \sigma_{\max}(S_i) := \sigma_{\max}(N_i). \quad (3.14)$$

Let β_i be the vector obtained from selecting the corresponding elements of $\hat{\alpha}_i$ in (3.12). Define $\tilde{J}_{AB,i}$ as the unique solution of

$$N_i \tilde{J}_{AB,i} = \beta_i. \quad (3.15)$$

Let 2^I denote the power set of any countable set I. Define the maps

$$h_1 : \left(\bigcup_{i\in I} \operatorname{int}(S_i)\right)^\complement \to 2^I, \quad x \mapsto \{i \in I : x \in \partial S_i\} \quad (3.16a)$$

$$h_2 : 2^I \to I, \qquad J \mapsto j^* := \operatorname{argmax}\left\{ \left\| \tilde{J}_{AB,j} \right\|_2 : j \in J \right\} \quad (3.16b)$$

$$h : \left(\bigcup_{i\in I} \operatorname{int}(S_i)\right)^\complement \to \mathbb{R}^d, \quad x \mapsto h_2 \circ h_1(x). \quad (3.16c)$$

We then define the *discrete probability current* $\tilde{J}_{AB} : S \to \mathbb{R}^d$ associated to a Voronoi tesselation $\{S_i\}_{i\in I}$ via the function

$$\tilde{J}_{AB}(x) := \begin{cases} \tilde{J}_{AB,i} & x \in \operatorname{int}(S_i) \\ \tilde{J}_{AB,h(x)} & \text{otherwise.} \end{cases} \quad (3.17)$$

That is, the function $\tilde{J}_{AB}$ is piecewise constant on the interiors of the Voronoi cells, and for any x not in the interior of a Voronoi cell, the probability current is the L^2-norm maximising current chosen from among the cells whose boundary contains x. Since the complement of the interiors consist of points lying on the interfaces between two or more Voronoi cells, we define the function to be the average of the corresponding interiors. For the purposes of establishing a convergence of $\tilde{J}_{AB}$ to J_{AB} in the $L^2(S,\mu;\mathbb{R}^d)$ topology the definition of $\tilde{J}_{AB}$ on complement of the interiors is not important, as this set has Lebesgue measure zero. However, for the purposes of defining discrete streamlines – which we consider in §3.4 – this case is important.

Note that the specification of N_i via (3.13) could be avoided by simply specifying N_i to be any square matrix with linearly independent rows chosen from M_i. We have chosen (3.13) in order to optimise the constant in the error bound of Theorem 3.16. To

Theorem 3.16 (Error bound in discrete probability current). *Let J_{AB} be a globally Lipschitz function on S with Lipschitz constant L, and let $\{S_i\}_{i\in I}$ be a Voronoi tesselation of S with width ρ. Then for every $i \in I$, there exists some $C_i > 0$ that does not depend on ρ, such that*

$$\left\| \tilde{J}_{AB,i} - J_{AB}(x) \right\|_2 \leq \rho C_i, \quad \forall x \in S_i.$$

In particular, for any discretised probability current $\tilde{J}_{AB} : S \to \mathbb{R}^d$ that is equal to $\tilde{J}_{AB,i}$ on $\operatorname{int}(S_i)$ *for every $i \in I$, we have*

$$\left\| \tilde{J}_{AB} - J_{AB} \right\|_{L^2(S,\mu;\mathbb{R}^d)} \leq \rho \max_{i\in I} C_i. \tag{3.18}$$

Proof. The second statement follows from the first, the union of the facets in any tesselation has μ-measure zero, and since μ is a probability measure on S. Thus it suffices to prove the first statement. Fix an arbitrary $x \in S_i$. Since N_i is invertible, we may use (3.14) to write

$$\begin{aligned}
\left\| \tilde{J}_{AB,i} - J_{AB}(x) \right\|_2 &= \left\| N_i^{-1} N_i \left(\tilde{J}_{AB,i} - J_{AB}(x) \right) \right\|_2 \\
&\leq \sigma_{\max}\left(N_i^{-1}\right) \left\| N_i \left(\tilde{J}_{AB,i} - J_{AB}(x) \right) \right\|_2 \\
&= \frac{1}{\sigma_{\min}(N_i)} \left\| N_i \left(\tilde{J}_{AB,i} - J_{AB}(x) \right) \right\|_2 \\
&= \frac{1}{\sigma_{\min}(S_i)} \left\| N_i \left(\tilde{J}_{AB,i} - J_{AB}(x) \right) \right\|_2 .
\end{aligned}$$

By relabelling the indices in $\mathcal{N}_i$, we may assume without loss of generality that the rows of N_i are given by the outer unit normals $(n_{ij})_{j=1}^d$ to S_i; this implies that the corresponding entries of β_i are given by $(\alpha_{i,j})_{j=1}^d$. Therefore, it follows from (3.15) that

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

Using (3.11), the Cauchy-Schwarz inequality, the fact that $\|n_{ij}\|_2 = 1$ for all $i \in I$ and $j \in \mathcal{N}_i$, 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

$$\begin{aligned}
\left| \alpha_{i,j} - n_{ij} \cdot J_{AB}(x) \right| &= \left| \frac{1}{\sigma(\partial S_i \cap \partial S_j)} \int_{\partial S_i \cap \partial S_j} n_{ij} \cdot (J_{AB}(y) - J_{AB}(x)) \, \mathrm{d}\sigma(y) \right| \\
&\leq \frac{1}{\sigma(\partial S_i \cap \partial S_j)} \int_{\partial S_i \cap \partial S_j} \|n_{ij}\|_2 \, \|J_{AB}(y) - J_{AB}(x)\|_2 \, \mathrm{d}\sigma(y) \\
&\leq \frac{1}{\sigma(\partial S_i \cap \partial S_j)} \int_{\partial S_i \cap \partial S_j} L \, \|y - x\|_2 \, \mathrm{d}\sigma(y) \\
&\leq \frac{L\rho}{\sigma(\partial S_i \cap \partial S_j)} \sigma(\partial S_i \cap \partial S_j) = L\rho.
\end{aligned}$$

Combining the preceding inequalities yields

$$\left\| \tilde{J}_{AB,i} - J_{AB}(x) \right\|_2 \leq \frac{\sqrt{d} L \rho}{\sigma_{\min}(S_i)},$$

so that the desired error bound holds with $C_i := \sqrt{d} L \sigma_{\min}^{-1}(S_i)$. □

Define the *smallest singular value of a Voronoi tesselation* $\{S_i\}_{i \in I}$ via

$$\sigma_{\min}(\{S_i\}_{i \in I}) := \min_{i \in I} \sigma_{\min}(S_i). \tag{3.19}$$

The error bound (3.18) of Theorem 3.16 can then be written as

$$\left\| \tilde{J}_{AB} - J_{AB} \right\|_{L^2(S,\mu;\mathbb{R}^d)} \leq \sqrt{d} L \left(\sigma_{\min}(\{S_i\}_{i \in I}) \right)^{-1} \rho. \tag{3.20}$$

To obtain a convergence theorem from the bound above, we must ensure that the sequence $(\sigma_{\min}(\{S_i^{(n)}\}_{i \in I(n)}))_{n \in \mathbb{N}}$ associated to a sequence of Voronoi tesselations with decreasing widths remains bounded away from zero. We will accomplish this in Corollary 3.19 below.

Define the *inner radius* of a set $A \subset \mathbb{R}^d$,

$$\operatorname{inrad}(A) := \sup\{r > 0 \ : \ \exists y \in S_i \text{ such that } B(y,r) \subset S_i\}, \tag{3.21}$$

where $B(y,r)$ is the open d-dimensional Euclidean ball with centre y and radius r. We shall use the *degeneracy ratio* of a d-dimensional set A,

$$\delta(A) := \frac{\operatorname{diam}(A)}{\operatorname{inrad}(A)}, \tag{3.22}$$

to quantify how close a d-dimensional set A is to being $(d-1)$-dimensional. Note that $\delta(\cdot)$ is invariant under homotheties, since the scaling factor appears in both the numerator and denominator. We motivate the degeneracy ratio as follows: Suppose that A and B are d-dimensional sets so that $\delta(A) > \delta(B)$, and let A' be a homothetic image of A such that $\operatorname{diam}(A') = \operatorname{diam}(B)$ and $\delta(A') = \delta(A)$; then

$$\frac{\operatorname{diam}(A')}{\operatorname{inrad}(A')} = \frac{\operatorname{diam}(A)}{\operatorname{inrad}(A)} > \frac{\operatorname{diam}(B)}{\operatorname{inrad}(B)},$$

and since $\operatorname{diam}(A') = \operatorname{diam}(B)$, it holds that A' and B are of equal 'size' in the sense that they are contained in balls of the same radius $\operatorname{diam}(B) = \operatorname{diam}(A')$, but A' is 'thinner' or closer to being $(d-1)$-dimensional, because it also holds that $\operatorname{inrad}(A') < \operatorname{inrad}(B)$.

The next lemma establishes a relationship between the degeneracy ratio of a polytope S and its smallest singular value.

Lemma 3.17. *Let S_1 and S_2 be d-dimensional polytopes in $\mathbb{R}^d$, and let N_1 and N_2 be the corresponding matrices constructed according to (3.13). It holds that*

$$\sigma_{\min}(S_1) < \sigma_{\min}(S_2) \Rightarrow \delta(S_1) < \delta(S_2).$$

Proof. Recall that in the singular value decomposition $M = U\Sigma V^*$ of a matrix M, the singular values have the interpretation of scaling factors along the directions specified by the unitary matrix V of right singular vectors. In particular, as the smallest singular value $\sigma_{\min}(M)$ approaches zero, the image of the unit ball under M becomes 'flatter', i.e. the extent of the ellipsoid along the minor axis approaches zero, and hence the ellipsoid becomes closer to a $(d-1)$-dimensional object. Therefore, the inequality $\sigma_{\min}(S_1) < \sigma_{\min}(S_2)$ implies that the image of the unit ball under N_1 is closer to being $(d-1)$-dimensional than the corresponding image under N_2. By definition of $\delta(\cdot)$, it follows that $\delta(S_1) < \delta(S_2)$. □

Next, we define the *inner width* of a Voronoi tesselation $\{S_i\}_{i\in I}$ by

$$\rho_{\text{in}}(\{S_i\}_{i\in I}) := \min_{i\in I} \operatorname{inrad}(S_i). \tag{3.23}$$

By comparing (2.2) and (3.23), it follows that $\rho_{\text{in}}(\{S_i\}_{i\in I}) \leq \rho(\{S_i\}_{i\in I})$. Note that equality is equivalent to $\operatorname{inrad}(S_i) = \operatorname{diam}(S_i)$ for every $i \in I$, in which case every S_i is a ball of the same radius. Since no polytope is a ball, it follows that the inequality above is strict for Voronoi tesselations.

In order to state our convergence theorem, we will need the following proposition.

Proposition 3.18. *Let $(\rho_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in $(0,\infty)$ decreasing to zero. For each n, let $\{S_i^{(n)}\}_{i\in I(n)}$ be a Voronoi tesselation of width ρ_n. Suppose that there exists some $0 < c < 1$ that depends on $(\rho_n)_{n\in\mathbb{N}}$ but not on n, such that the tesselations satisfy the* uniformity condition

$$c\rho_n = \rho(\{S^{(n)}_{i\in I(n)}) \leq \rho_{\text{in}}(\{S^{(n)}\}_{i\in I(n)}). \tag{UC}$$

Then there exists some $c' > 0$ that does not depend on n such that

$$c' < \sigma_{\min}(\{S_i^{(n)}\}_{i\in I(n)}), \quad \forall n \in \mathbb{N},$$

i.e. the smallest singular values of the tesselations are bounded away from zero.

Proof. By the uniformity condition (UC) and the definitions (2.2) and (3.23) of $\rho(\cdot)$ and $\rho_{\text{in}}(\cdot)$, it holds that for all $n \in \mathbb{N}$,

$$\delta(S_j^{(n)}) = \frac{\operatorname{diam}(S_j)}{\operatorname{inrad}(S_j)} \leq \frac{\operatorname{diam}(S_j)}{\min_{i\in I(n)} \operatorname{inrad}(S_i^{(n)})} \leq \frac{\max_{i\in I(n)} \operatorname{diam}(S_i^{(n)})}{\min_{i\in I(n)} \operatorname{inrad}(S_i^{(n)})} = \frac{\rho(\{S_i^{(n)}\}_{i\in I(n)})}{\rho_{\text{in}}(\{S_i^{(n)}\}_{i\in I(n)})} \leq \frac{1}{c}$$

for all $j \in I(n)$. In particular, this implies that the collection $\{\delta(S_j^{(n)}) : j \in I(n), n \in \mathbb{N}\}$ is bounded from above. Thus, by Lemma 3.17, the collection $\{\sigma_{\min}(S_j^{(n)}) : j \in I(n), n \in \mathbb{N}\}$ is bounded from below, and by Corollary A.2, it follows that the latter collection is bounded away from zero, as desired. □

We now state our convergence theorem for the discrete probability current. Recall that $\rho(\{S_i\}_{i\in I}$ denotes the width of a tesselation $\{S_i\}_{i\in I}$; see (2.2).

Corollary 3.19 (Convergence of discrete probability current under uniformity condition)**.** *Let $(\rho_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in $(0,\infty)$ decreasing to zero. For each n, let $\{S_i^{(n)}\}_{i\in I(n)}$ be a Voronoi tesselation of width ρ_n, and suppose that there exists some $c > 0$ that does not depend on n such that the tesselations satisfy the uniformity condition (UC). Then the sequence $(\tilde{J}_{AB}^{(n)})_{n\in\mathbb{N}}$ of corresponding discrete probability currents satisfies*

$$\left\| \tilde{J}_{AB}^{(n)} - J_{AB} \right\|_{L^2(S,\mu;\mathbb{R}^d)} \leq C\sqrt{d}L\rho_n,$$

for some $C > 0$ that does not depend on n.

Proof. Let $(\rho_n)_{n\in\mathbb{N}}$ be given, and fix an arbitrary $n \in \mathbb{N}$. By Theorem 3.16, we have (3.20):

$$\left\| \tilde{J}_{AB}^{(n)} - J_{AB} \right\|_2 \leq \sqrt{d}L \left(\sigma_{\min}(\{S_i^{(n)}\}_{i\in I(n)})\right)^{-1} \rho_n, \quad \forall n \in \mathbb{N}.$$

Since (UC) holds, Proposition 3.18 implies that

$$\left(\sigma_{\min}(\{S_i^{(n)}\}_{i\in I(n)})\right)^{-1} \leq (c')^{-1}$$

for some $c' > 0$ that does not depend on n. □

3.4. Streamlines

We now define discrete streamlines, using the discrete probability current defined in the previous section, and prove that the discrete streamlines converge to the streamlines of the diffusion process. Recall that, in TPT, a streamline between the reactant set A and the product set B for a given initial condition $s_0 \in \partial A$ is the solution $(s(t))_{t\in[0,T(s_0)]}$ of the initial value problem

$$s(0) = s_0, \quad \frac{\mathrm{d}s}{\mathrm{d}t}(t') = J_{AB}(s(t')), \quad t' \in [0, T(s_0)] \tag{3.24a}$$

$$T(s_0) := \inf\{t' > 0 \ : \ s(t') \in B\}. \tag{3.24b}$$

In addition to Assumption 3.1, we will make the following assumption.

Assumption 3.20. *For all $s_0 \in \partial A$, $T(s_0)$ is finite and strictly positive.*

The idea of our construction of the discrete streamline $(\tilde{s}(t'))_{t'\geq 0}$ will be to use the discrete probability current $\tilde{J}_{AB} : S \to \mathbb{R}^d$ from (3.17). Since $\tilde{J}_{AB}$ is piecewise constant on the interiors of the cells of a Voronoi tesselation, we would like to use (3.24) to construct for every initial condition $\tilde{s}_0$ a streamline that is *piecewise linear* on the interiors of the Voronoi cells. Since the discrete probability current converges to the true probability current in the L^2-topology, we expect that for any given initial condition $s_0 \in \partial A$, the discrete streamline starting at s_0 converges to the true streamline starting at s_0.

A problem with the approach described above is that the discrete probability current $\tilde{J}_{AB}$ is not continuous. Hence, the standard existence and uniqueness theorems for solutions of ordinary differential equations do not apply. Nevertheless, we can define a continuous, piecewise linear trajectory that solves the analogue of (3.24a) with J_{AB} replaced by $\tilde{J}_{AB}$: since (3.24a) is equivalent to

$$s(t') = s_0 + \int_0^{t'} J_{AB}(s(r))\mathrm{d}r, \quad 0 \leq t' \leq T(s_0), \tag{3.25}$$

we may define the *discrete streamline* by

$$\tilde{s}(t') = \tilde{s}_0 + \int_0^{t'} \tilde{J}_{AB}(\tilde{s}(r))\mathrm{d}r, \quad 0 \leq t' \leq T(\tilde{s}_0). \tag{3.26}$$

Note that if $\tilde{s}_0 = s_0$, then the upper time limit in (3.26) equals $T(s_0)$. Furthermore, since $\tilde{J}_{AB}$ is piecewise constant on the interiors of the Voronoi cells, it follows immediately from (3.26) that the discrete streamline has the desired properties.

Theorem 3.21 (Error bounds for discrete streamlines)**.** *Let $J_{AB} : S \to \mathbb{R}^d$ be a globally Lipschitz vector field with Lipschitz constant L, and let $\{S_i\}_{i\in I}$ be a Voronoi tesselation of S with width $\rho > 0$. Fix $s_0 = \tilde{s}_0 \in \partial A$, and let s and $\tilde{s}$ be the true and discrete streamlines defined according to (3.25) and (3.26) respectively. Then*

$$\|s - \tilde{s}\|_{L^2([0,T(s_0)])} \leq C\rho$$

where $C := \sqrt{dT(s_0)}L \max_{i\in I} \sigma_{\min}^{-1}(S_i) \exp(LT(s_0))$ does not depend on ρ.

Proof. Since $s_0 = \tilde{s}_0$ we have from (3.25) and (3.26) that

$$\|s(t) - \tilde{s}(t)\|_2 \leq \int_0^t \left\| J_{AB}(s(r)) - \tilde{J}_{AB}(\tilde{s}(r))\right\|_2 \mathrm{d}r.$$

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

$$\begin{aligned}\left\|J_{AB}(s(r)) - \tilde{J}_{AB}(\tilde{s}(r))\right\|_2 &\leq \|J_{AB}(s(r)) - J_{AB}(\tilde{s}(r))\|_2 + \left\|J_{AB}(\tilde{s}(r)) - \tilde{J}_{AB}(\tilde{s}(r))\right\|_2 \\ &\leq L\,\|s(r) - \tilde{s}(r)\|_2 + \left\|J_{AB}(\tilde{s}(r)) - \tilde{J}_{AB}(\tilde{s}(r))\right\|_2 .\end{aligned}$$

Recall that Theorem 3.16 and its proof yield

$$\left\|J_{AB}(x) - \tilde{J}_{AB,i}(x)\right\|_2 \leq \rho\frac{\sqrt{d}L}{\sigma_{\min}(S_i)}, \quad \forall x \in S_i.$$

Therefore,

$$\left\|J_{AB}(\tilde{s}(r)) - \tilde{J}_{AB}(\tilde{s}(r))\right\|_2 \leq \rho\sqrt{d}L \max_{i\in I} \sigma_{\min}^{-1}(S_i), \quad \forall r \in [0, T(s_0)].$$

Combining the preceding estimates yields

$$\|s(t) - \tilde{s}(t)\|_2 \leq \rho\sqrt{d}L \max_{i\in I} \sigma_{\min}^{-1}(S_i) + L\int_0^t \|s(r) - \tilde{s}(r)\|_2 \,\mathrm{d}r.$$

By the Gronwall-Bellman inequality, it follows that

$$\|s(t) - \tilde{s}(t)\|_2 \leq \rho\sqrt{d}L \max_{i\in I} \sigma_{\min}^{-1}(S_i) \exp\left(Lt\right), \quad \forall t \in [0, T(s_0)],$$

which proves that

$$\|s(t) - \tilde{s}(t)\|_2 \leq \rho\sqrt{d}L \max_{i\in I} \sigma_{\min}^{-1}(S_i) \exp\left(LT(s_0)\right), \quad \forall t \in [0, T(s_0)].$$

Since the right-hand side of the inequality above does not depend on t, the desired conclusion follows. □

The final step is to prove convergence of the discrete streamlines, given a sequence of Voronoi tesselations with decreasing widths.

Corollary 3.22 (Convergence of discrete streamlines)**.** *Let $(\rho_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in $(0, \infty)$ decreasing to zero. For each n, let $\{S_i^{(n)}\}_{i\in I(n)}$ be a Voronoi tesselation of width ρ_n, such that for some $0 < c < 1$ that does not depend on n, the uniformity condition (UC) holds. Fix $s_0 = \tilde{s}_0 \in \partial A$, let s be the streamline generated by s, and let $\tilde{s}_{AB}^{(n)}$ be the discrete streamline generated by $\tilde{s}_0$ and the discrete probability current associated to $\{S_i^{(n)}\}_{i\in I(n)}$. Then there exists some $C > 0$ that does not depend on n, such that*

$$\left\|\tilde{s}^{(n)} - s\right\|_2 \leq C\rho_n.$$

Proof. By Theorem 3.21, we have

$$\left\|s - \tilde{s}^{(n)}\right\|_{L^2([0,T(s_0)])} \leq \sqrt{dT(s_0)}L \exp(LT(s_0)) \max_{i\in I} \sigma_{\min}^{-1}(S_i^{(n)})\rho_n.$$

Since the uniformity condition (UC) holds, Proposition 3.18 implies that $\{\sigma_{\min}(S_i^{(n)}) : i \in I(n), n \in \mathbb{N}\}$ is bounded away from zero by some scalar that does not depend on n. This completes the proof. □

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 the Freie Universität Berlin within the Excellence Initiative of the German Research Foundation (DFG).

References

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

[2] Nicholas Boekelheide, Romelia Salomón-Ferrer, and Thomas F. Miller. Dynamics and dissipation in enzyme catalysis. *Proceedings of the National Academy of Sciences*, 2011.

[3] Gregory R Bowman, Vincent A Voelz, and Vijay S Pande. Taming the complexity of protein folding. *Current Opinion in Structural Biology*, 21(1):4 – 11, 2011.

[4] Maria Cameron and Eric 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] Q. Du, V. Faber, and M. Gunzburger. Centroidal voronoi tessellations: Applications and algorithms. *SIAM Review*, 41(4):637–676, 1999.

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

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

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

[9] H. 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.

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

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

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

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

[14] Frank Noé, Christof Schütte, Eric Vanden-Eijnden, Lothar Reich, and Thomas 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.

[15] Nuria Plattner and Frank Noé. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and markov models. *Nature Communications*, 6(7653), 2015.

[16] Jan-Hendrik Prinz, Bettina Keller, and Frank Noé. Probing molecular kinetics with markov models: metastable states, transition pathways and spectroscopic observables. *Phys. Chem. Chem. Phys.*, 13, 2011.

[17] Daniel-Adriano Silva, Gregory R. Bowman, Alejandro Sosa-Peinado, and Xuhui 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.

[18] Alain-Sol Sznitman. *Brownian motion, obstacles and random media*. Springer Science & Business Media, 1998.

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

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

[21] Eric Vanden-Eijnden and Maddalena Venturoli. Markovian milestoning with voronoi tessellations. *The Journal of Chemical Physics*, 130(19):194101, 2009.

[22] Eric Vanden-Eijnden, Maddalena Venturoli, Giovanni Ciccotti, and Ron Elber. On the assumptions underlying milestoning. *The Journal of Chemical Physics*, 129(17):174102, 2008.

A. Basic facts about polytopes

$$\sigma_{\min}(A)\|x-y\|_2 \leq \|A(x-y)\|_2 \leq \sigma_{max}(A)\|x-y\|_2 \tag{A.1}$$

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 polytopes in $\mathbb{R}^2$ of full dimension with the smallest number of facets are 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. □

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 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. □