

**Weierstraß-Institut  
für Angewandte Analysis und Stochastik  
Leibniz-Institut im Forschungsverbund Berlin e. V.**

Preprint

ISSN 2198-5855

**Estimation of the infinitesimal generator  
by square-root approximation**

Luca Donati<sup>1</sup> , Martin Heida<sup>2</sup> , Marcus Weber<sup>3</sup> , Bettina Keller<sup>1</sup>

submitted: August 8, 2017

<sup>1</sup> FU Berlin, Physical and Theoretical Chemistry  
Takustr. 3  
14195 Berlin  
Germany  
E-Mail: ldonati@zedat.fu-berlin.de  
E-Mail: bettina.keller@fu-berlin.de

<sup>2</sup> Weierstraß-Institut  
Mohrenstr. 39  
10117 Berlin  
Germany  
E-Mail: martin.heida@wias-berlin.de

<sup>3</sup> Zuse Institute Berlin  
Takustr. 7  
14195 Berlin  
Germany  
E-Mail: weber@zib.de

No. 2416  
Berlin 2017



---

2008 *Physics and Astronomy Classification Scheme*. 02.50.Ga, 05.10.Gg.

*Key words and phrases.* Molecular simulation, Markov state models, transfer operator, molecular kinetics.

This research has been funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 *Scaling Cascades in Complex Systems*, Projects A05 *Probing scales in equilibrated systems by optimal nonequilibrium forcing*, B05 *Origin of the scaling cascades in protein dynamics* and C05 *Effective models for interfaces with many scales*.

Edited by  
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)  
Leibniz-Institut im Forschungsverbund Berlin e. V.  
Mohrenstraße 39  
10117 Berlin  
Germany

Fax: +49 30 20372-303  
E-Mail: preprint@wias-berlin.de  
World Wide Web: <http://www.wias-berlin.de/>

# Estimation of the infinitesimal generator by square-root approximation

Luca Donati, Martin Heida, Marcus Weber, Bettina Keller

## Abstract

For the analysis of molecular processes, the estimation of time-scales, i.e., transition rates, is very important. Estimating the transition rates between molecular conformations is – from a mathematical point of view – an invariant subspace projection problem. A certain infinitesimal generator acting on function space is projected to a low-dimensional rate matrix. This projection can be performed in two steps. First, the infinitesimal generator is discretized, then the invariant subspace is approximated and used for the subspace projection. In our approach, the discretization will be based on a Voronoi tessellation of the conformational space. We will show that the discretized infinitesimal generator can simply be approximated by the geometric average of the Boltzmann weights of the Voronoi cells. Thus, there is a direct correlation between the potential energy surface of molecular structures and the transition rates of conformational changes. We present results for a 2d-diffusion process and Alanine dipeptide.

## 1 Introduction

Molecular processes, like conformational changes or binding processes, have been modeled as time-continuous processes switching between a finite number of molecular states [1]. In continuous space, the transition behavior can be described by a propagator[19]. By Galerkin’s method, this propagator is projected to a finite dimensional matrix. A wide collection of tools[17] have been developed to numerically estimate the discretized version of the propagator, i.e., the transition probability matrix which acts on small subsets of the (discretized) conformational space. This transition matrix and the corresponding Markov chain have a unique stationary distribution, as long as the process is ergodic and aperiodic. Associated to the propagator, there exists another operator, named infinitesimal generator[19, 16]. The infinitesimal generator provides transition pattern in terms of rates in conformational space. The discretized version of the generator is, thus, called rate matrix. For estimating this rate matrix, we will apply the square root approximation in the following. This approximation has been described in literature before[14, 18]. It had been conjectured, that this approximation provides a meaningful rate matrix. Just recently[7] it has been proved, indeed, that a rate matrix built on that approximation and on an arbitrary tessellation of the space, like a Voronoi tessellation, converges to the backward generator of the Smoluchowski equation, that describes the dynamics of a molecular system in the limit of high friction. In this paper, we give a detailed explanation of the mentioned operators that describe the dynamics of a molecular system. We recall the main steps to prove that the square root approximation converges to an infinitesimal

generator in the limit of small Voronoi cells. This approximation is valid except for one unknown scalar factor – the flux of the system. Finally, we discuss an approach for determining this flux. In our case, we use the PCCA+ method to build the rate matrix of the conformations (metastable states) and scale the transition rates properly. We will present results for a two dimensional system and for Alanine-dipeptide.

## 2 Theory

To fully explain the theoretical background, we need to introduce a set of related operators (fig. 2). In section 2.1, we discuss the relation between the transfer operators (top face of the blocks in fig. 2-A and 2.B), and in section 2.2, we discuss the corresponding infinitesimal generators (bottom face of the blocks in fig. 2-A and 2.B).

### 2.1 Propagator, Koopman operator, and transfer operator

Let  $\{\mathbf{x}_t\} \in \Gamma$  be a dynamic process defined on a state space  $\Gamma$ . A large collection of these processes (or equivalently: systems) at time  $t$  is called an ensemble, and the distribution of the processes across  $\Gamma$  at time  $t$  is given by the probability density function  $\rho_t(\mathbf{x}) : \Gamma \rightarrow \mathbb{R}_{\geq 0}$ . Assuming that the process is *Markovian*, we can introduce a continuous operator, called *propagator*[19, 17] (or *Perron-Frobenius operator*)  $\mathcal{P}(\tau) : L^1(\Gamma) \rightarrow L^1(\Gamma)$  such that

$$\rho_{t+\tau}(\mathbf{y}) = \mathcal{P}(\tau)\rho_t(\mathbf{y}) = \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau)\rho_t(\mathbf{x})d\mathbf{x} \quad (1)$$

where  $p(\mathbf{x}, \mathbf{y}; \tau)$  is the transition probability density, i.e. the conditional probability to find the process in state  $\mathbf{y}$  after a *lag time*  $\tau$ , given the initial state  $\mathbf{x}$ . Thus, the operator  $\mathcal{P}(\tau)$  propagates probability densities  $\rho_t(\mathbf{x})$  forward in time.

The *backward transfer operator* [16, 19, 8] (or *Koopman operator*)  $\mathcal{K}(\tau) : L^\infty(\Gamma) \rightarrow L^\infty(\Gamma)$  is defined as

$$\mathbb{E}_{\mathbf{x}}[f(\mathbf{x}_t)] = \mathcal{K}(\tau)f(\mathbf{x}) = \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau)f(\mathbf{y})d\mathbf{y} \quad (2)$$

where the function  $f : \Gamma \rightarrow \mathbb{R}$  is an observable of the system, and  $\mathbb{E}_{\mathbf{x}}[f(\mathbf{x}_t)]$  is the expected value of this observable at time  $t + \tau$ , given that the process has been started in  $\mathbf{x}_t = \mathbf{x}$  at time  $t$ . Thus, the initial point  $\mathbf{x}$  is fixed, and the integral in eq. (2) is computed over the arrival points  $\mathbf{y}$ .

The Koopman operator  $\mathcal{K}(\tau)$  is the left-adjoint of the propagator  $\mathcal{P}(\tau)$

$$\begin{aligned} \langle \mathcal{K}(\tau)f, \rho_t \rangle &= \int_{\Gamma} \left[ \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau)f(\mathbf{y})d\mathbf{y} \right] \rho_t(\mathbf{x})d\mathbf{x} \\ &= \int_{\Gamma} f(\mathbf{y}) \left[ \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau)\rho_t(\mathbf{x})d\mathbf{x} \right] d\mathbf{y} \\ &= \langle f, \mathcal{P}(\tau)\rho_t \rangle. \end{aligned} \quad (3)$$

In fig. 2, the adjointness is visualized by the edges between the red and the blue face. Note that the expression in eq. 3 can be interpreted as the time evolution of the ensemble

average of  $f$ , given that the ensemble was initially distributed according to  $\rho_t(x)$  at time  $t$

$$\begin{aligned}\mathbb{E}[f(\mathbf{x}_t)] &= \int_{\Gamma} f(\mathbf{y}) \rho_{t+\tau}(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Gamma} f(\mathbf{y}) \left[ \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau) \rho_t(\mathbf{x}) d\mathbf{x} \right] d\mathbf{y} \\ &= \langle f, \mathcal{P}(\tau) \rho_t \rangle = \langle \mathcal{K}(\tau) f, \rho_t \rangle.\end{aligned}\quad (4)$$

If the process is *ergodic* and *time-homogeneous* (i.e. if there are no time-dependent external forces), then there exists a unique invariant measure  $\pi(\mathbf{x})$  such that for all lag times  $\tau$ :

$$\pi(\mathbf{y}) = \mathcal{P}(\tau) \pi(\mathbf{y}) = \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau) \pi(\mathbf{x}) d\mathbf{x}. \quad (5)$$

For a canonical ensemble (NVT), the invariant measure is given by the Boltzmann probability density:

$$\pi(\mathbf{x}) = \frac{\exp(-\beta \mathcal{H}(\mathbf{x}))}{Z}, \quad Z = \int_{\Gamma} \exp(-\beta \mathcal{H}(\mathbf{x})) d\mathbf{x}, \quad (6)$$

where  $\mathcal{H}(\mathbf{x})$  is the classical Hamiltonian of the system,  $\beta = \frac{1}{k_B T}$ ,  $T$  is the temperature and  $k_B$  the Boltzmann constant.

Using the invariant measure, we define the *forward transfer operator* [17]  $\mathcal{T}(\tau) : L_{\pi}^1(\Gamma) \rightarrow L_{\pi}^1(\Gamma)$ :

$$u_{t+\tau}(\mathbf{y}) = \mathcal{T}(\tau) u_t(\mathbf{y}) = \frac{1}{\pi(\mathbf{y})} \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau) u_t(\mathbf{x}) \pi(\mathbf{x}) d\mathbf{x} \quad (7)$$

which propagates relative probability densities  $u_t(\mathbf{x})$  forward in time, where  $u_{t+\tau}(\mathbf{y}) = \rho_t(\mathbf{y})/\pi(\mathbf{y})$ . Operators which act on probability densities are represented on the left face in fig. 2-A, while operators which act on weighted probability densities are represented on the right face.

For *reversible* processes, i.e. if

$$\pi(\mathbf{x}) p(\mathbf{x}, \mathbf{y}; \tau) = \pi(\mathbf{y}) p(\mathbf{y}, \mathbf{x}; \tau), \quad (8)$$

the forward transfer operator and the Koopman operator are self-adjoint with respect to the weighted scalar product

$$\langle u, v \rangle_{\pi} = \int_{\Gamma} u(\mathbf{x}) \pi(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}, \quad (9)$$

which can be seen by

$$\begin{aligned}\langle u, \mathcal{T}(\tau) v \rangle_{\pi} &= \int_{\Gamma} u(\mathbf{y}) \pi(\mathbf{y}) \left[ \frac{1}{\pi(\mathbf{y})} \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau) \pi(\mathbf{x}) v(\mathbf{x}) d\mathbf{x} \right] d\mathbf{y} \\ &= \int_{\Gamma} \int_{\Gamma} u(\mathbf{y}) \pi(\mathbf{x}) p(\mathbf{x}, \mathbf{y}; \tau) v(\mathbf{x}) d\mathbf{x} d\mathbf{y} \\ &= \int_{\Gamma} \int_{\Gamma} u(\mathbf{y}) \pi(\mathbf{y}) p(\mathbf{y}, \mathbf{x}; \tau) v(\mathbf{x}) d\mathbf{x} d\mathbf{y} \\ &= \int_{\Gamma} \left[ \frac{1}{\pi(\mathbf{x})} \int_{\Gamma} p(\mathbf{y}, \mathbf{x}; \tau) \pi(\mathbf{y}) u(\mathbf{y}) d\mathbf{y} \right] \pi(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}\end{aligned}$$

$$= \langle \mathcal{T}(\tau)u, v \rangle_\pi. \quad (10)$$

Note that an analogous calculation holds for Koopman operator  $\mathcal{K}(\tau)$ . The most important insight is that for reversible processes we find  $\mathcal{K}(\tau) = \mathcal{T}(\tau)$  i.e.

$$\frac{1}{\pi(\mathbf{y})} \int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau) \pi(\mathbf{x}) v(\mathbf{x}) d\mathbf{x} = \int_{\Gamma} p(\mathbf{y}, \mathbf{x}; \tau) v(\mathbf{x}) d\mathbf{x}. \quad (11)$$

Thus, for reversible processes, the right face in fig. 2-A. "collapses" into a single line (fig. 2-B). This implies that, for reversible processes, the transfer operator is the adjoint of the propagator with respect to the Euclidean scalar product

$$\langle u, \mathcal{P}(\tau)v \rangle = \langle \mathcal{K}(\tau)u, v \rangle = \langle \mathcal{T}(\tau)u, v \rangle. \quad (12)$$

The expression in eq. 12 can be interpreted as the correlation function  $\text{cor}(u, v; \tau)$  between  $u(\mathbf{x})$  and  $v(\mathbf{x})$ .

## 2.2 Evolution equation and the generator

The time-derivative of the probability density at time  $t = 0$  is then given as

$$\begin{aligned} \dot{\rho}_t(\mathbf{x})|_{t=0} &= \left. \frac{\partial}{\partial t} \rho_t(\mathbf{x}) \right|_{t=0} \\ &= \lim_{\tau \rightarrow 0} \frac{\rho_\tau(\mathbf{x}) - \rho_0(\mathbf{x})}{\tau} \\ &= \lim_{\tau \rightarrow 0} \frac{\mathcal{P}(\tau)\rho_0(\mathbf{x}) - \mathcal{P}(0)\rho_0(\mathbf{x})}{\tau} \\ &= \lim_{\tau \rightarrow 0} \frac{\mathcal{P}(\tau) - \mathcal{P}(0)}{\tau} \rho_0(\mathbf{x}) \\ &= \mathcal{L}\rho_0(\mathbf{x}). \end{aligned} \quad (13)$$

where  $\mathcal{L}$  is the infinitesimal generator of  $\mathcal{P}(\tau)$ .  $\mathcal{L}$  is a function of the time-derivative of the transition probability density

$$\begin{aligned} \mathcal{L}\rho_0(\mathbf{x}) &= \int_{\Gamma} \lim_{\tau \rightarrow 0} \frac{p(\mathbf{x}, \mathbf{y}; \tau) - p(\mathbf{x}, \mathbf{y}; 0)}{\tau} \rho_0(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Gamma} \left. \frac{\partial}{\partial \tau} p(\mathbf{x}, \mathbf{y}; \tau) \right|_{\tau=0} \rho_0(\mathbf{x}) d\mathbf{x} \\ &= \left. \frac{\partial}{\partial \tau} \mathcal{P}(\tau) \right|_{\tau=0} \rho_0(\mathbf{x}), \end{aligned} \quad (14)$$

and the time evolution equation is

$$\dot{\rho}_t(\mathbf{x}) = \mathcal{L}\rho_t(\mathbf{x}). \quad (15)$$

If the dynamic process  $\{\mathbf{x}_t\}$  evolves according to the Langevin equation or the Brownian equation of motion, eq. 14 represents the Fokker-Planck equation. If the process evolves according to Hamiltonian dynamics,  $\mathcal{L}$  represents the classical Liouville operator. The relation between an operator and its infinitesimal generator is represented by vertical lines in fig. 2.

The left-adjoint of  $\mathcal{L}$  is  $\mathcal{Q}$ , the *infinitesimal generator*  $\mathcal{K}(\tau)$

$$\begin{aligned}\langle f, \mathcal{L}\rho_0 \rangle &= \int_{\Gamma} f(y) \left[ \int_{\Gamma} \frac{\partial}{\partial \tau} p(\mathbf{x}, \mathbf{y}; \tau) \Big|_{\tau=0} \rho_0(\mathbf{x}) d\mathbf{x} \right] d\mathbf{y} \\ &= \int_{\Gamma} \left[ \int_{\Gamma} \frac{\partial}{\partial \tau} p(\mathbf{x}, \mathbf{y}; \tau) \Big|_{\tau=0} f(y) d\mathbf{y} \right] \rho_0(\mathbf{x}) d\mathbf{x} \\ &= \langle \mathcal{Q}f, \rho_0 \rangle\end{aligned}\tag{16}$$

Analogous to the relation between  $\mathcal{L}$  and the propagator  $\mathcal{P}(\tau)$  (eq. 13 and eq. 14), the infinitesimal generator  $\mathcal{Q}$  is linked to the Koopman operator  $\mathcal{K}(\tau)$  by

$$\mathcal{Q}f(\mathbf{x}) = \left. \frac{\partial \mathcal{K}(\tau)}{\partial \tau} \right|_{\tau=0} f(\mathbf{x})\tag{17}$$

$$= \lim_{\tau \downarrow 0} \frac{\mathcal{K}(\tau)f(\mathbf{x}) - \mathcal{K}(0)f(\mathbf{x})}{\tau}\tag{18}$$

$$= \lim_{\tau \downarrow 0} \frac{\mathbb{E}_{\mathbf{x}}[f(\mathbf{x}_t)] - f(\mathbf{x})}{\tau}\tag{19}$$

(20)

Thus, we can derive  $\mathcal{Q}$  by defining the adjoint of  $\mathcal{L}$ , or by regarding the time-derivative of  $\mathbb{E}_{\mathbf{x}}[f(\mathbf{x}_t)]$ . Note that this definition of the infinitesimal generator is consistent with the usual definition of the generator of a diffusion process as [19, 8, 23, 11, 16].

In an analogous way, we can derive the infinitesimal generator  $\mathcal{L}_{\pi}$  of  $\mathcal{T}(\tau)$  by examining the time-derivative of  $u_t(\mathbf{x})$ . However, for reversible processes, the self-adjointness of  $\mathcal{T}(\tau)$  and  $\mathcal{K}(\tau)$  with respect to eq. (1), extends to their infinitesimal generators (fig. 2-B)

$$\mathcal{Q}f(\mathbf{x}) = \left. \frac{\partial \mathcal{T}(\tau)}{\partial \tau} \right|_{\tau=0} f(\mathbf{x})\tag{21}$$

$$= \lim_{\tau \downarrow 0} \frac{\mathcal{T}(\tau)f(\mathbf{x}) - \mathcal{T}(0)f(\mathbf{x})}{\tau}\tag{22}$$

$$= \mathcal{L}_{\pi}f(\mathbf{x})\tag{23}$$

Note, the term infinitesimal generator  $\mathcal{L}$  means that the semigroup of  $\mathcal{P}(\tau)$  with  $\tau > 0$  can be generated by

$$\mathcal{P}(\tau) = \exp(\mathcal{L}\tau)\tag{24}$$

Analogous relation holds for all other operators/infinitesimal generator pairs in this manuscript.

### 2.3 Discretization of the time evolution operator and the generator

Consider an arbitrary partition of the state space in Voronoi cells  $\Gamma = \cup_{i=1}^n \Omega_i$  with associated indicator functions given by

$$\mathbf{1}_i(\mathbf{x}) := \begin{cases} 1, & \mathbf{x} \in \Omega_i \\ 0, & \mathbf{x} \notin \Omega_i \end{cases}\tag{25}$$

Assume that  $\rho = \sum_i \rho_i \mathbf{1}_i$  is a piecewise constant function solving (15) and for given  $\varphi \in \mathbb{R}^n$  multiply (15) with  $\varphi := \sum_i \varphi_i \pi \mathbf{1}_i$ . We obtain

$$\begin{aligned} \sum_i \langle \dot{\rho}_i \mathbf{1}_i, \varphi_i \pi \mathbf{1}_i \rangle &= \langle \mathcal{Q}^* \rho, \sum_i \varphi_i \mathbf{1}_i \rangle_\pi \\ &= \langle \rho, \sum_i \varphi_i \mathcal{Q} \mathbf{1}_i \rangle_\pi \\ &= \sum_i \sum_j \varphi_i \rho_j \langle \mathbf{1}_j, \mathcal{Q} \mathbf{1}_i \rangle_\pi. \end{aligned} \quad (26)$$

Hence, we obtain

$$\forall \varphi \in \mathbb{R}^n \quad \sum_i \dot{\rho}_i \varphi_i = \sum_i \sum_j \varphi_i \rho_j Q_{ij} \quad \Leftrightarrow \quad \dot{\rho} = \rho^T Q, \quad (27)$$

where

$$Q_{ij} = \frac{\langle \mathbf{1}_i, \mathcal{Q} \mathbf{1}_j \rangle_\pi}{\langle \mathbf{1}_i, \mathbf{1}_i \rangle_\pi}. \quad (28)$$

By Galerkin discretization, we can construct also a discretization matrix  $\mathbf{T}(\tau)$  of the operator  $\mathcal{T}$  via

$$T_{ij}(\tau) = \frac{\langle \mathcal{T}(\tau) \mathbf{1}_i, \mathbf{1}_j \rangle_\pi}{\langle \mathbf{1}_i, \mathbf{1}_i \rangle_\pi} = \frac{\langle \mathbf{1}_i, \mathcal{T}(\tau) \mathbf{1}_j \rangle_\pi}{\langle \mathbf{1}_i, \mathbf{1}_i \rangle_\pi}. \quad (29)$$

The entries  $T_{ij}(\tau)$  describe the transition probability from a cell  $\Omega_i$  to  $\Omega_j$  after a time  $\tau$ :

$$T_{ij}(\tau) = \mathbb{P}[\mathbf{x}(t + \tau) \in \Omega_j | \mathbf{x}(t) \in \Omega_i] \quad (30)$$

Several methods have been developed to estimate  $\mathbf{T}(\tau)$ , e.g. by Markov State Models[17].

The matrix  $\mathbf{Q}$  is a valid discretization of the operators  $\mathcal{Q}$  and  $\mathcal{L} = \mathcal{Q}^*$ , while the matrix  $\mathbf{T}(\tau)$  is a discretization of the operators  $\mathcal{T}(\tau)$  and  $\mathcal{T}(\tau) = \mathcal{T}(\tau)^*$ . As one can see from (26) and (27), when applied from the left, the matrices  $\mathbf{Q}$  and  $\mathbf{T}(\tau)$  act on the weighted space, when applied from the right, act on the euclidean space.

The elements  $Q_{ij}$ , are the time-derivative of the transition probabilities  $T_{ij}(\tau)$  (eq. 29 and 30) evaluated at  $\tau = 0$

$$Q_{ij} = \frac{\langle \mathbf{1}_i, \frac{\partial \mathcal{T}(\tau)}{\partial \tau} \mathbf{1}_j \rangle_\pi}{\langle \mathbf{1}_i, \mathbf{1}_i \rangle_\pi} = \frac{\partial}{\partial \tau} \frac{\langle \mathbf{1}_i, \mathcal{T}(\tau) \mathbf{1}_j \rangle_\pi}{\langle \mathbf{1}_i, \mathbf{1}_i \rangle_\pi} \Big|_{\tau=0} = \frac{\partial T_{ij}(\tau)}{\partial \tau} \Big|_{\tau=0} \quad (31)$$

Intuitively, the elements of the diagonal  $Q_{ii}$  describe the transition rate out the states  $i$ ; while the elements  $Q_{ij}/Q_{ii}$  describe the transition rate from the cell  $\Omega_i$  to  $\Omega_j$ . For infinitesimal small values of the lag time  $\tau$ , the processes which start in cell  $\Omega_i$  at time  $t$ , can only reach directly neighboring (adjacent) cells by time  $t + \tau$ . Thus for infinitesimal small  $\tau$ ,  $T_{ij}(\tau)$  and  $Q_{ij}$  are zero if  $\Omega_i$  and  $\Omega_j$  are non-adjacent cells.

## 2.4 Gauss theorem

We will use the Gauss theorem as stated in Ref. [23] to derive a formulation of the rate matrix  $\mathbf{Q}$  that can be estimated numerically.

**Gauss theorem.** Given a Voronoi tessellation of the state space  $\Gamma = \cup_{i=1}^n \Omega_i$  and the discretized transfer operator  $\mathbf{T}(\tau)$ , the matrix  $\mathbf{Q} := \frac{\partial \mathbf{T}(\tau)}{\partial \tau} \Big|_{\tau=0}$  satisfies

$$Q_{ij} = \frac{1}{\pi_i} \oint_{\partial\Omega_i \partial\Omega_j} \Phi(\mathbf{z}) \pi(\mathbf{z}) dS(\mathbf{z}) \quad (32)$$

where  $\partial\Omega_i \partial\Omega_j$  is the common surface between the cell  $\Omega_i$  and  $\Omega_j$ ,  $\pi_i$  is the Boltzmann density of the cell  $i$  and  $\Phi(\mathbf{z})$  denotes the flux of the configurations  $\mathbf{z} \in \partial\Omega_i \partial\Omega_j$ , through the infinitesimal surface  $\partial\Omega_i \partial\Omega_j$ .

**Proof** The transition probability density  $p(\mathbf{x}, \mathbf{y}; \tau)$  describes the probability that the system will visit the state  $\mathbf{y}$ , starting in  $\mathbf{x}$ , after a time  $\tau$ . Because the system will be always in some state, it yields  $\int_{\Gamma} p(\mathbf{x}, \mathbf{y}; \tau) d\mathbf{y} = 1$ . Thus the conservation of the probability density can be associated to the mass conservation of a fluid, that moves in the state space  $\Gamma$ , transporting the properties of the system. If we consider an ensemble  $S_{\tau, \mathbf{x}, n}$  of  $n \rightarrow \infty$  trajectories, starting in the state  $\mathbf{x}$ , of length  $\tau$ , then we can think to the time evolution of the ensemble, like the time evolution of a continuum (fluid); where each trajectory (each particle of the fluid) will end in a different state  $\mathbf{y}$  according to  $p(\mathbf{x}, \mathbf{y}; \tau)$ . We write the continuity equation:

$$\frac{\partial p(\mathbf{x}, \mathbf{y}; \tau)}{\partial \tau} \Big|_{\tau=0} = -\nabla_{\mathbf{y}} \cdot \mathbf{j} \quad (33)$$

where  $\mathbf{j} = p(\mathbf{x}, \mathbf{y}; \tau) \mathbf{v}(\mathbf{x})$  is the density flux and  $\mathbf{v}(\mathbf{x})$  is the flow velocity. We interpret the flux of the probability density as the probability per unit area per unit time, that a trajectory of the ensemble (particle of the fluid) passes through a surface. While the flow velocity vector is a vector field that represents the velocity with which the system moves from the state  $\mathbf{x}$  to  $\mathbf{y}$  (or the velocity of the fluid that moves from  $\mathbf{x}$  to  $\mathbf{y}$ ). We now use the continuity equation, to rewrite  $\mathbf{Q}$ :

$$\begin{aligned} Q_{ij} &:= \frac{\partial T_{ij}(\tau)}{\partial \tau} \Big|_{\tau=0} \\ (a) &= \frac{\partial}{\partial \tau} \left[ \frac{1}{\pi_i} \int_{\Omega_j} \int_{\Omega_i} p(\mathbf{x}, \mathbf{y}; \tau) \pi(\mathbf{x}) d\mathbf{x} d\mathbf{y} \right] \Big|_{\tau=0} \\ &= \frac{1}{\pi_i} \int_{\Omega_j} \int_{\Omega_i} \frac{\partial p(\mathbf{x}, \mathbf{y}; \tau)}{\partial \tau} \Big|_{\tau=0} \pi(\mathbf{x}) d\mathbf{x} d\mathbf{y} \\ (b) &= \frac{1}{\pi_i} \int_{\Omega_j} \int_{\Omega_i} -\nabla_{\mathbf{y}} \cdot \mathbf{j}|_{\tau=0} \pi(\mathbf{x}) d\mathbf{x} d\mathbf{y} \\ (c) &= \frac{1}{\pi_i} \int_{\Omega_j} \oint_{\partial\Omega_i} -\mathbf{j}|_{\tau=0} \cdot \mathbf{n}_i \pi(\mathbf{x}) dS(\mathbf{x}) d\mathbf{y} \\ (d) &= \frac{1}{\pi_i} \int_{\Omega_j} \oint_{\partial\Omega_i} -\delta_{\mathbf{x}=\mathbf{y}} \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}_i \pi(\mathbf{x}) dS(\mathbf{x}) d\mathbf{y} \\ (e) &= \frac{1}{\pi_i} \oint_{\partial\Omega_i \partial\Omega_j} \Phi(\mathbf{z}) \pi(\mathbf{z}) dS(\mathbf{z}) \end{aligned} \quad (34)$$

where we have used:

- (a) The result of the Galerkin discretization of the transfer operator in eq. (29).

(b) The continuity equation.

(c) The divergence theorem. The vector  $\mathbf{n}_i$  is the unit vector normal to the surface  $\partial\Omega_i$ .

(d) Because  $\tau = 0$ , we have

$$p(\mathbf{x}, \mathbf{y}; \tau) = \delta_{\mathbf{x}=\mathbf{y}} \quad (35)$$

Then the density flux is  $\mathbf{j}|_{\tau=0} = \delta_{\mathbf{x}=\mathbf{y}} \mathbf{v}(\mathbf{x})$ .

(e) If  $\tau = 0$  only instantaneous transitions between neighbor cells have to be taken in account. Thus, the only points that satisfy  $\mathbf{x} = \mathbf{y}$ , are the points on the intersecting surface  $\partial\Omega_i \partial\Omega_j$ . We denote those points, in the integral (e), with  $\mathbf{z}$ . The quantity  $\Phi(\mathbf{z})$  denotes the flux of the configurations  $\mathbf{z}$  trough the infinitesimal surface  $\partial\Omega_i$ . Note that  $\Phi(\mathbf{z}) = -\delta_{\mathbf{x}=\mathbf{y}} \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}_i = \delta_{\mathbf{x}=\mathbf{y}} \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}_j$  where  $\mathbf{n}_j$  is the unit vector normal to the surface  $\partial\Omega_j$ .

□

## 2.5 Square Root Approximation

Following [14], we now show how the matrix (32) can be approximated by Square Root Approximation (SqRA). First of all, we rewrite eq.32, multiplying and dividing by the quantity

$$s_{ij} = \oint_{\partial\Omega_i \partial\Omega_j} \pi(\mathbf{z}) dS(\mathbf{z}) \quad (36)$$

that represents the Boltzmann density of the intersecting surface  $\partial\Omega_i \partial\Omega_j$ . Thus  $Q_{ij}$  becomes

$$\begin{aligned} Q_{ij} &= \frac{s_{ij}}{\pi_i} \oint_{\partial\Omega_i \partial\Omega_j} \Phi(\mathbf{z}) \frac{\pi(\mathbf{z})}{s_{ij}} dS(\mathbf{z}) \\ &= \frac{s_{ij}}{\pi_i} \langle \Phi \rangle_{ij} \end{aligned} \quad (37)$$

the quantity  $\langle \Phi \rangle_{ij}$  represents the mean value of the flux, trough  $\partial\Omega_i \partial\Omega_j$ , if the potential energy function was flat ( $V(\cdot) = 0$ ), weighted by the Boltzmann density of the intersecting surface  $\partial\Omega_i \partial\Omega_j$ .

In what follows, we make the following assumptions

- There exists a constant  $\hat{\Phi}$  such that for all  $i, j$  it holds  $\hat{\Phi} = \langle \Phi \rangle_{ij}$ .
- The cells are so small such that  $\pi$  is almost constant on every  $\Omega_i$  and on every interface  $\partial\Omega_i \partial\Omega_j$ , i.e.  $\pi|_{\Omega_i} \approx \pi_i$  and  $\pi|_{\partial\Omega_i \partial\Omega_j} \approx \pi_{ij}$ .

Due to assumption 1, the matrix  $\mathbf{Q}$  can be rewritten as

$$Q'_{ij} = \frac{s_{ij}}{\pi_i} \hat{\Phi} \quad (38)$$

The quantity  $s_{ij}$ , i.e. the Boltzmann density of the intersecting surface between two neighbor cells  $\Omega_i$  and  $\Omega_j$ , is a surface integral. Due to assumption,  $s_{ij}$  can be approximated by the Boltzmann density  $\pi_{ij}$  through

$$\begin{aligned} s_{ij} &= \oint_{\partial\Omega_i \partial\Omega_j} \pi(\mathbf{z}) dS(\mathbf{z}) \\ &\approx \pi_{ij} \end{aligned} \quad (39)$$

and

$$Q''_{ij} = \frac{\pi(\hat{\mathbf{z}})}{\pi_i} \hat{\Phi} \quad (40)$$

Since all points on the surface  $\partial\Omega_i \partial\Omega_j$  are equidistant from the centers of the cells  $\Omega_i$  and  $\Omega_j$ , we estimate  $\pi_{ij}$  by the geometric average

$$\pi_{ij} = \sqrt{\pi_i \pi_j}. \quad (41)$$

Thus, the rate matrix is given as

$$Q'''_{ij} = \frac{\sqrt{\pi_i \pi_j}}{\pi_i} \hat{\Phi} = \sqrt{\frac{\pi_j}{\pi_i}} \hat{\Phi} \quad (42)$$

If we consider a molecular system in the limit of high friction, the dynamics is governed by the Smoluchowski equation and the stationary distribution depends only on the potential energy function  $V(\cdot)$ . In this case, the Boltzmann density of the midpoint  $\hat{\mathbf{z}}$  can be easily estimated as

$$\pi(\hat{\mathbf{z}}) = \sqrt{\pi_i \pi_j} = \frac{\exp\{-\beta \frac{1}{2} [V(\Omega_i) + V(\Omega_j)]\}}{Z} \quad (43)$$

The diagonal entries of  $\mathbf{Q}$  are  $Q_{ii} = -\sum_{j \neq i} Q_{ij}$ , thus the sum of the rows is zero  $\sum_j Q_{ij} = 0$ . The matrix  $\mathbf{Q}$  has also the property that the ordinary differential equation (27) reads

$$\left. \frac{\partial \rho_i}{\partial \tau} \right|_{\tau=0} = C \sum_{i \sim j} (\rho_j Q_{ji} - \rho_i Q_{ij}) \quad (44)$$

where  $C$  is a normalization constant and the notation  $i \sim j$  denote neighboring cells.

## 2.6 Convergence of the rate matrix

The square root approximation of the infinitesimal generator  $\mathcal{Q}$  is based on a Voronoi tessellation of the state space. In principle, it is not clear if this kind of approximation has a physically meaningful limit, whenever the number of Voronoi cells tend to infinity. If this limit exists, it is furthermore not clear whether the limit operator is physically reasonable. An arbitrary refinement strategy of increasing the number of Voronoi cells will probably not converge. However, a recent mathematical study [7] which will be published separately, shows that under suitable assumptions on the Voronoi tessellation, the square root approximation converges towards the backward generator of the Smoluchowski equation, i.e., towards the Langevin dynamics for the limit of high friction.

In what follows, let  $\varepsilon$  denote the maximal diameter of the cells and for given  $\varepsilon > 0$  let  $(P_i^\varepsilon)_i$  be the set of points that generate the Voronoi tessellation and let  $\Omega_i^\varepsilon$  be the Voronoi cell that corresponds to  $P_i^\varepsilon$ . For any continuous function  $u$ , we write  $\rho_i^\varepsilon := \rho(P_i^\varepsilon)$ . In

particular, for the function  $\pi(x) := \exp(-\beta V(x))$  we write  $v_i^\varepsilon := \sqrt{\pi_i^\varepsilon}$  and using (40) we denote the right hand side of (44) as

$$(\mathcal{L}^\varepsilon \rho)_i^\varepsilon := C_\varepsilon \sum_{i \sim j} \left( \rho_j^\varepsilon \frac{v_i^\varepsilon}{v_j^\varepsilon} - \rho_i^\varepsilon \frac{v_j^\varepsilon}{v_i^\varepsilon} \right), \quad (45)$$

where  $\sum_{i \sim j}$  relates to the sum over all neighboring cells  $\Omega_j^\varepsilon$  of the cell  $\Omega_i^\varepsilon$  and where we interpret  $(\mathcal{L}^\varepsilon \rho)_i^\varepsilon$  as a function which is constant on every  $\Omega_i^\varepsilon$ . Written in a formal way, for the scaling  $C_\varepsilon = \varepsilon^{-2}$  and a suitable positive definite symmetric matrix  $A_{\text{hom}} \in \mathbb{R}^{n \times n}$  it holds for twice continuously differentiable functions  $u$  that[7]

$$C_\varepsilon \sum_{i \sim j} \left( \rho_j^\varepsilon \frac{v_i^\varepsilon}{v_j^\varepsilon} - \rho_i^\varepsilon \frac{v_j^\varepsilon}{v_i^\varepsilon} \right) \rightarrow \nabla \cdot (A_{\text{hom}} \nabla \rho(x)) + \beta \nabla \cdot (\rho(x) A_{\text{hom}} \nabla V(x)) \quad \text{as } \varepsilon \rightarrow 0 \quad (46)$$

provided that  $P_i^\varepsilon \rightarrow x$  as  $\varepsilon \rightarrow 0$ . In case that the Voronoi tessellations are isotropic, we find  $A_{\text{hom}} = a\mathbb{I}$ , where  $a > 0$  is a constant and  $\mathbb{I}$  is the identity. However, if for some reason the tessellations are systematically anisotropic, the right hand side of (46) can be brought into the form of a classical Fokker-Planck operator via a coordinate transform. Hence, we also call the right hand side of (46) a Fokker-Planck operator. In what follows, we will roughly explain how the convergence (46) can be obtained.

There exists a simpler version of (45) known as *the discrete Laplace operator*  $\mathcal{F}u$  having the form

$$(\mathcal{F}^\varepsilon \rho)_i^\varepsilon := C_\varepsilon \sum_{i \sim j} (\rho_j^\varepsilon - \rho_i^\varepsilon).$$

Note that the use of  $\mathcal{L}$  and  $\mathcal{F}$  is opposite in[7]. It turned out that the understanding of the asymptotic behavior of  $\mathcal{F}^\varepsilon$  is essential for the study of the asymptotic behavior of  $\mathcal{L}^\varepsilon$ .

In case the point process is a rectangular grid (Fig. 3), the operator  $\mathcal{F}^\varepsilon$  has been studied intensively and in broad generality from both physicists (as generator of a markovian process that models Brownian motion, see the review [4]) and mathematicians (rigorous results, see the review [3]). The notion of discrete Laplace operator can be understood as follows: On the lattice  $\varepsilon\mathbb{Z}^n$  (that consists of all points  $x \in \mathbb{R}^n$  such that  $(\varepsilon^{-1}x) \in \mathbb{Z}^n$ , see Figure 3), the *discrete derivative* in the  $j$ -th direction is given by  $d_{j,\varepsilon}\rho(x) := \frac{1}{\varepsilon}(\rho(x + \varepsilon e_j) - \rho(x))$ , where  $e_j$  is the  $j$ -th unit vector. The *second order discrete derivative* is given by

$$\begin{aligned} d_j^2 \rho(x) &:= \frac{1}{\varepsilon} (d_j \rho(x) - d_j \rho(x - \varepsilon e_j)) \\ &= \frac{1}{\varepsilon^2} [(\rho(x + \varepsilon e_j) - \rho(x)) + (\rho(x - \varepsilon e_j) - \rho(x))]. \end{aligned} \quad (47)$$

Hence, we obtain for  $x_i^\varepsilon \in \varepsilon\mathbb{Z}^n$  that

$$\sum_{j=1}^n d_j^2 \rho(x_i^\varepsilon) = \varepsilon^{-2} \sum_{j \sim i} (\rho(x_j^\varepsilon) - \rho(x_i^\varepsilon)) = (\mathcal{F}^\varepsilon \rho)_i^\varepsilon, \quad (48)$$

where now  $i \sim j$  relates to all neighbors  $x_j^\varepsilon \in \varepsilon\mathbb{Z}^n$  of  $x_i^\varepsilon$  s.t.  $|x_j^\varepsilon - x_i^\varepsilon| = \varepsilon$ . We now show that for twice continuously differentiable functions  $u$  it holds

$$(\mathcal{F}^\varepsilon \rho)_i^\varepsilon \rightarrow \Delta \rho \quad \text{as } \varepsilon \rightarrow 0 \text{ if } x_i^\varepsilon \rightarrow x. \quad (49)$$

In order to show this, we use Taylor's formula, i.e.

$$\rho(x + \varepsilon e_i) - \rho(x) = \partial_i \rho(x) \varepsilon + \frac{1}{2} \partial_i^2 \rho(x) \varepsilon^2 + \sum_{k=3}^{\infty} \frac{1}{k!} \partial_i^k \rho(x) \varepsilon^k \quad (50)$$

and hence

$$\begin{aligned} d_j^2 \rho(x_i^\varepsilon) &= \frac{1}{\varepsilon^2} \left( \partial_j \rho(x_i^\varepsilon) \varepsilon + \frac{1}{2} \partial_j^2 \rho(x_i^\varepsilon) \varepsilon^2 + \sum_{k=3}^{\infty} \frac{1}{k!} \partial_j^k \rho(x_i^\varepsilon) \varepsilon^k \right) \\ &\quad + \frac{1}{\varepsilon^2} \left( -\partial_j \rho(x_i^\varepsilon) \varepsilon + \frac{1}{2} \partial_j^2 \rho(x_i^\varepsilon) \varepsilon^2 + \sum_{k=3}^{\infty} \frac{1}{k!} \partial_j^k \rho(x_i^\varepsilon) (-\varepsilon)^k \right) \\ &= \partial_j^2 \rho(x_i^\varepsilon) + \mathcal{O}(\varepsilon). \end{aligned} \quad (51)$$

From this, we obtain that (49) holds.

We will now use the above insights to formally understand the asymptotic behavior of the operator

$$(\mathcal{L}^\varepsilon \rho)_i := \varepsilon^{-2} \sum_{i \sim j} \left( \rho_j^\varepsilon \frac{v_i^\varepsilon}{v_j^\varepsilon} - \rho_i^\varepsilon \frac{v_j^\varepsilon}{v_i^\varepsilon} \right) \quad \text{where } v_i^\varepsilon = \exp \left( -\frac{1}{2} \beta V(x_i^\varepsilon) \right) \quad (52)$$

on the lattice  $\varepsilon \mathbb{Z}^n$ . Writing  $V_i^\varepsilon := V(x_i^\varepsilon)$  and using the Taylor formula we obtain

$$\frac{v_i^\varepsilon}{v_j^\varepsilon} = 1 - \frac{1}{2} \beta (V_i^\varepsilon - V_j^\varepsilon) + \frac{1}{8} \beta^2 (V_i^\varepsilon - V_j^\varepsilon)^2 - \mathcal{O}((V_i^\varepsilon - V_j^\varepsilon)^3) \quad (53)$$

and inserting this expansion into (52) we obtain

$$\begin{aligned} (\mathcal{L}^\varepsilon \rho)_i &= \varepsilon^{-2} \sum_{i \sim j} \left( (\rho_j^\varepsilon - \rho_i^\varepsilon) + \frac{\beta}{2} (\rho_j^\varepsilon + \rho_i^\varepsilon) (V_j^\varepsilon - V_i^\varepsilon) \right) \\ &\quad + \varepsilon^{-2} \sum_{i \sim j} \left( \frac{1}{8} \beta^2 (V_i^\varepsilon - V_j^\varepsilon)^2 + \mathcal{O}((V_i^\varepsilon - V_j^\varepsilon)^4) \right) (\rho_j^\varepsilon - \rho_i^\varepsilon) \\ &\quad + \varepsilon^{-2} \sum_{i \sim j} (\rho_j^\varepsilon + \rho_i^\varepsilon) \mathcal{O}((V_i^\varepsilon - V_j^\varepsilon)^3). \end{aligned} \quad (54)$$

We know that for small  $\varepsilon \approx 0$  it holds  $(V_i^\varepsilon - V_j^\varepsilon) \approx \varepsilon \nabla V(x_i^\varepsilon) + \mathcal{O}(\varepsilon^2)$  and equivalently  $(\rho_j^\varepsilon - \rho_i^\varepsilon) \approx \varepsilon \nabla \rho(x_i^\varepsilon) + \mathcal{O}(\varepsilon^2)$ . Therefore, we obtain

$$(\mathcal{L}^\varepsilon \rho)_i^\varepsilon = \varepsilon^{-2} \sum_{i \sim j} \left( (\rho_j^\varepsilon - \rho_i^\varepsilon) + \frac{\beta}{2} (\rho_j^\varepsilon + \rho_i^\varepsilon) (V_j^\varepsilon - V_i^\varepsilon) \right) + \mathcal{O}(\varepsilon) \quad (55)$$

and using once more the Taylor expansion for  $u$  and  $V$  in (55) we further obtain

$$\begin{aligned} (\mathcal{L}^\varepsilon \rho)_i^\varepsilon &= \varepsilon^{-2} \sum_{i \sim j} (\rho_j^\varepsilon - \rho_i^\varepsilon) + \varepsilon^{-2} \rho_i^\varepsilon \sum_{i \sim j} \beta (V_j^\varepsilon - V_i^\varepsilon) + \varepsilon^{-2} \sum_{i \sim j} \frac{\beta}{2} (\rho_j^\varepsilon - \rho_i^\varepsilon) (V_j^\varepsilon - V_i^\varepsilon) + \mathcal{O}(\varepsilon) \\ &= \varepsilon^{-2} \sum_{i \sim j} (\rho_j^\varepsilon - \rho_i^\varepsilon) + \varepsilon^{-2} \rho_i^\varepsilon \sum_{i \sim j} \beta (V_j^\varepsilon - V_i^\varepsilon) + \sum_{i \sim j} \frac{\beta}{2} \partial_j \rho(x_i^\varepsilon) \cdot \partial_j V(x_i^\varepsilon) + \mathcal{O}(\varepsilon). \end{aligned} \quad (56)$$

Thus, as  $\varepsilon \rightarrow 0$  we observe that

$$\begin{aligned} (\mathcal{L}^\varepsilon \rho)_i^\varepsilon &\rightarrow \Delta \rho(x) + \beta \rho(x) \Delta V(x) + \beta \nabla \rho(x) \cdot \nabla V(x) & x_i^\varepsilon \rightarrow x \\ &= \Delta \rho(x) + \beta \nabla \cdot (\rho(x) \nabla V(x)) \end{aligned}$$

on the Grid  $\varepsilon \mathbb{Z}^n$ . Hence, we recover (46) with  $A_{\text{hom}} = 1$  for the cubic Voronoi tessellation. On arbitrary Voronoi tessellations, things become more involved. In particular, the convergence (49) or calculations like in (51) and (56) do not hold any more, as they rely on the rectangular structure of  $\mathbb{Z}^n$ . However, the key ideas of the proof remain the same with the difference that some terms which explicitly cancel out in the above calculation only vanish in a “statistically averaged” sense, using G-convergence.

G-convergence is a concept from early stage in the development of Homogenization theory and is extremely rarely used (refer to [9]), since other concepts are usually much better suited. In the discrete setting, G-convergence can be formulated in the following sense: The operator  $\mathcal{F}^\varepsilon$  is called G-convergent if there exists a symmetric positive definite matrix  $A_{\text{hom}}$  such that for every continuous  $f : \Omega \rightarrow \mathbb{R}$  the sequence  $u^\varepsilon$  of solutions to

$$-(\mathcal{F}^\varepsilon \rho^\varepsilon)_i := - \sum_{j \sim i} (\rho_j^\varepsilon - \rho_i^\varepsilon) = f(P_i^\varepsilon) \quad (57)$$

converges in  $L^2(\Omega)$  to the solutions  $\rho$  of  $-\nabla \cdot (A_{\text{hom}} \nabla \rho) = f$ , where we interpret  $\rho^\varepsilon$  as a function that is constant on every cell  $\Omega_i^\varepsilon$ . Hence, G-convergence and convergence of the SQRA-operator are more or less equivalent conditions on the tessellation. In recent years, G-convergence (or (49)) has been proved for random operators

$$(\mathcal{F}_\omega^\varepsilon \rho)(x_i) := \frac{1}{\varepsilon^2} \sum_{j \sim i} \omega_{ij} (\rho(x_j) - \rho(x_i)) \quad (58)$$

on the grid  $\varepsilon \mathbb{Z}^n$  for a broad range of random coefficients  $\omega_{ij}$ , see the overview in [3]. However, for stationary and ergodic tessellations, the recent results in [2, 7] seem to be the only ones.

In conclusion we showed that the convergence (46) holds on the rectangular grid for sufficiently smooth functions  $u$ . The calculations suggest that the result also holds on more general grids. However, on such more general grids, the mathematics behind the convergence (46) becomes much more involved and is thus shifted to the article [7]. The results there are though more general as they state that solutions of  $\mathcal{L}^\varepsilon \rho^\varepsilon = f^\varepsilon$  converge to solutions of  $\mathcal{L}\rho = f$  provided  $f^\varepsilon \rightarrow f$  in  $L^2(\mathbb{Q})$ .

## 3 Methods

### 3.1 Discretization error

The Galerkin discretization of the continuous operators  $\mathcal{T}(\tau)$  and  $\mathcal{Q}$  described in the theory section is one source of errors. With regard to the transition probability matrix  $\mathbf{T}(\tau)$ , the discretization of the space causes the loss of information and the loss of the Markov property [17, 23, 13]. Let's consider a discretization of the space into  $n$  disjoint sets  $A_1, \dots, A_n$ , and assume that at time  $t > 0$ , the state  $\mathbf{x}(t)$  is in a certain set  $A_i$ .

The transition probability to the next set at time  $t + \tau$  depends on the position  $\mathbf{x}(t)$  inside the set  $A_i$ . Although the probability to be in state  $\mathbf{x}(t)$  at time  $t$  only depends on the previous position  $\mathbf{x}(t - \tau)$ , this Markov property is lost, on the level of the sets  $A_i$ . The transition behavior between the sets  $A_i$  is not Markovian in general. The deviation from Markovianity of the matrix  $\mathbf{T}(\tau)$  can be reduced, by choosing a proper partition of the state space. Alternatively, it can be proved that for a large enough value of  $\tau$ , the implied time scales  $t_i = -\frac{\tau}{\log \lambda_i}$  (where  $\lambda_i$  are the eigenvalues of the transition probability matrix  $\mathbf{T}(\tau)$ ) converge to constant values as a consequence of the Chapman-Kolmogorov equation.

With regard to the matrix  $\mathbf{Q}$ , two problems arise:

- If  $\tau$  is small, the matrix  $\mathbf{T}(\tau)$  is not Markovian, and the matrix  $\mathbf{Q}$  cannot be considered as a generator of  $\mathbf{T}(\tau)$ .
- If  $\tau$  is big enough to guarantee the Markovianity of  $\mathbf{T}(\tau)$ , then the generator  $\mathbf{Q}$  is the correct generator, but it is not physically meaningful. A proper generator is defined for  $\tau \rightarrow 0$ , in other words for instantaneous transitions that occur between neighboring sets. If  $\tau$  is too big, then  $\mathbf{Q}$  would describe instantaneous transition rates between non-neighboring sets, that are not physical if we are considering a time-continuous dynamics.

In conclusion, the matrix  $\mathbf{Q}$  is the Galerkin discretization of the infinitesimal generator  $\mathcal{Q}$ , but it is not the proper generator of the transition probability matrix  $\mathbf{T}(\tau)$ . In the next section, we will see that by using a different basis function, we can discretize the operators in transition matrices that satisfy respectively the Markov property and the generators properties.

### 3.2 Transition matrices in the space of conformations

We now consider a partition of the state-space into  $n$  disjoint Voronoi cells and a partition of the configuration space into  $n_c$  overlapping conformations, i.e., we introduce a membership matrix  $X \in \mathbb{R}^{n \times n_c}$  such that the entry  $X_{ij}$  provides the probability that a Voronoi cell  $\Omega_i$  belongs to a conformation  $C_j$ :

$$X_{ij} = X_j(\Omega_i) = \mathbf{P}[\Omega_i \in C_j] \in [0, 1], \quad i = 1, \dots, n; j = 1, \dots, n_c \quad (59)$$

The matrix  $X$  is nonnegative. Its entries form a partition of the unity, such that the row sum is  $\sum_{j=1}^{n_c} X_{ij} = 1 \quad \forall i$ . In this section, we use the notation  $X_{i\bullet}$  to denote row vectors of size  $n_c$  and  $X_{\bullet j}$  to denote column vectors of size  $n$ . Because of metastability of the conformations, a state  $\mathbf{x}$  tends to not leave its starting conformation  $C_j$ . The product between the rate matrix  $\mathbf{Q}$  and the column vector  $X_{\bullet j}$  is almost zero for all  $j$ .

$$\mathbf{Q}X_{\bullet j} \approx \mathbf{0} \quad (60)$$

where  $\mathbf{0}$  is the zero vector of size  $n$ . For the identification of the conformations it can be assumed, that the column vectors  $X_{\bullet j}$  span the same linear space as the leading first  $n_c$  right eigenvectors  $F_j$  of the matrix  $\mathbf{Q}$ , with eigenvalues  $\theta_j$  near zero [23]. For writing down the basis transform between these two linear spaces, we call  $F$  the matrix  $n \times n_c$  of the eigenvectors  $F_j$  using the same notation regarding column and row vectors.

We now define an invertible transformation matrix  $\mathbf{A}$  of size  $n_c \times n_c$ , such that  $X_{ij} = \sum_{k=1}^{n_c} F_{ik} A_{kj}$  (or  $X_{\bullet j} = FA_{\bullet j}$ ), or in matrix notation  $X = FA$ .

Because  $\mathbf{Q}F_{\bullet j} = \theta_j F_{\bullet j}$ , from eq. (60)

$$\mathbf{Q}(FA_{\bullet j}) = \theta_j(FA_{\bullet j}) \approx \mathbf{0} \quad (61)$$

Because the rows of the matrix  $X$  define the coordinates of  $n$  points in  $n_c$  dimensions that lie on the standard  $n_c$ -simplex, we can use Robust Perron Cluster Analysis (PCCA+) [5] to find the optimal transformation matrix  $\mathbf{A}$  and the matrix of the membership functions  $X$ . The transition rate matrix  $\mathbf{Q}^c$  of the conformations is the Galerkin discretization of the operator  $\mathcal{Q}$  on the basis of the membership functions  $\chi_j(\mathbf{x}) = \sum_{i=1}^n X_{ij} \mathbf{1}_{\Omega_i}(\mathbf{x})$ . Since the membership functions are not orthogonal, the matrix  $Q^c$  is the product of two matrices generated by the corresponding inner products:

$$\mathbf{Q}^c = (\langle \chi, \chi \rangle_\pi)^{-1} \langle \chi, \mathcal{Q} \chi \rangle_\pi \quad (62)$$

The following lemma[12, 22] implies that if  $\mathcal{Q}$  is the infinitesimal generator of  $\mathcal{T}(\tau)$ , then the matrix  $\mathbf{Q}^c$  is the generator of  $\mathbf{T}^c(\tau)$ .

**Lemma** If the  $n_c$  eigenfunctions  $\mathcal{F} = \{\mathcal{F}_1, \dots, \mathcal{F}_{n_c}\}$  of  $\mathcal{Q}$ , associated to eigenvalues  $\theta_j \approx 0$ , are  $\pi$ -orthonormal and  $\chi = \mathcal{F}\mathcal{A}$  is a regular basis transformation of these eigenfunctions, then

$$\text{if } \exp(\tau \mathcal{Q}) = \mathcal{T}(\tau) \Rightarrow \exp(\tau \mathbf{Q}^c) = \mathbf{T}^c(\tau)$$

where  $\mathcal{A}$  is a linear transformation corresponding to the matrix  $A$ .

*Proof.* The eigenfunctions  $\mathcal{F}_j$  of  $\mathcal{Q}$  are eigenfunctions also of  $\mathcal{T}(\tau)$  with eigenvalues  $\lambda_j(\tau) = \exp(\tau \theta_j)$  (semigroup property).

$$\begin{aligned} \mathcal{T}(\tau)\mathcal{F}_j &= \lambda_j(\tau)\mathcal{F}_j = \exp(\tau \theta_j)\mathcal{F}_j \\ \mathcal{Q}\mathcal{F}_j &= \theta_j\mathcal{F}_j \end{aligned} \quad (63)$$

By replacing  $\chi = \mathcal{A}^\top \mathcal{F}^\top$  and  $\chi = \mathcal{F}\mathcal{A}$  in eq.62, then we have

$$\begin{aligned} \mathbf{Q}^c &= \frac{\langle \chi, \mathcal{Q} \chi \rangle_\pi}{\langle \chi, \chi \rangle_\pi} \\ &= \frac{\mathcal{A}^\top \langle \mathcal{F}^\top, \mathcal{Q}\mathcal{F} \rangle_\pi \mathcal{A}}{\mathcal{A}^\top \langle \mathcal{F}^\top, \mathcal{F} \rangle_\pi \mathcal{A}} \\ &= \frac{\mathcal{A}^\top \langle \mathcal{F}^\top, \Theta\mathcal{F} \rangle_\pi \mathcal{A}}{\mathcal{A}^\top \mathcal{A}} \\ &= \mathcal{A}^\top \langle \mathcal{F}^\top, \mathcal{F} \rangle_\pi \Theta \mathcal{A} \\ &= \mathcal{A}^\top \Theta \mathcal{A} \end{aligned} \quad (64)$$

In the same way, it is possible to prove that  $\mathbf{T}^c(\tau) = \mathcal{A}^\top \Lambda(\tau) \mathcal{A}$ . The matrices  $\Theta$  and  $\Lambda(\tau)$  denote respectively the diagonal matrices  $n_c \times n_c$  of the eigenvalues  $\theta_j$  and  $\lambda(\tau)_j$ . By exploiting the relation in eq. (63), it yields

$$\exp(\tau \mathcal{A}^\top \Theta \mathcal{A}) = \exp(\tau \mathbf{Q}^c) = \mathbf{T}^c(\tau) = \mathcal{A}^\top \Lambda(\tau) \mathcal{A} \quad (65)$$

Then  $\mathbf{Q}^c$  is an infinitesimal generator of  $\mathbf{T}^c(\tau)$ .  $\square$

The above lemma also provides a method to estimate the transition rate matrix of the conformations. The linear transformation  $\mathcal{A}$  is unknown, however, by Robust Perron Cluster Analysis, it is possible to estimate the optimal matrix  $\mathbf{A}$  starting from the rate matrix  $Q$  obtained by Square Root Approximation.

## 4 Flux estimation

In the theory section 2.5, we derived the approximation (40) of the transition rate between two neighboring states  $i$  and  $j$ , i.e.  $Q'''_{ij} = \sqrt{\frac{\pi_j}{\pi_i}} \hat{\Phi}$ . The quantity  $\hat{\Phi}$  denotes the flux through the intersecting surface, a quantity that does not depend on the potential energy function or on the cells  $\Omega_i$  and  $\Omega_j$ . It is hard to estimate this quantity numerically, because of the curse of dimensionality. In this article, we propose a method that exploits the relation  $\exp(\tau \mathbf{Q}^c) = \mathbf{T}^c(\tau)$  to estimate the flux.

Given the time-discrete trajectory  $x_t$  of a dynamical system governed by a Hamiltonian function  $\mathcal{H}(x)$  that well samples the state space according to the Boltzmann distribution  $\pi(x) = \exp(-\beta \mathcal{H}(x))/Z$ . From this sampling, we discretize the state space with a Voronoi tessellation and we construct the matrix

$$\tilde{Q}_{ij} = \sqrt{\frac{\pi_j}{\pi_i}} = \frac{Q'''_{ij}}{\hat{\Phi}} \quad (66)$$

that differs from the rate matrix in equation (40) because we are neglecting the constant flux. We remark that  $i$  and  $j$  are the indices of neighbor cells. The same trajectory can also be used to construct a Markov State Model[17], a method to estimate the transition probability matrix  $\mathbf{T}(\tau)$ .

By PCCA+ we reduce both of the matrices  $\tilde{\mathbf{Q}}$  and  $\mathbf{T}(\tau)$ , to the transition matrices  $\tilde{\mathbf{Q}}^c$  and  $\mathbf{T}^c(\tau)$  between the conformations. Thus thewe obtain

$$T_{ij}^c(\tau) = \exp(\tau Q'''_{ij}^c) = \exp\left(\tau \hat{\Phi} \tilde{Q}_{ij}^c\right) \quad (67)$$

If  $\tilde{\theta}_i$  are the eigenvalues of  $\tilde{\mathbf{Q}}^c$  and  $\lambda_i(\tau)$  are the eigenvalues of  $\mathbf{T}^c(\tau)$ , the flux is given as

$$\hat{\Phi} = \frac{\log \lambda_i(\tau)}{\tau \tilde{\theta}_i} \quad \forall i \quad (68)$$

The flux should not depend on the choice of the eigenvalues. Nevertheless, in section 5 we will see, that depending on the quality of the Voronoi discretization, the fluxes obtained from different eigenvalues converge or not to the same value. This fact is useful to evaluate the quality of the rate matrix.

We remark that the constant  $\hat{\Phi}$  found in eq. (68), from a mathematical point of view is the same constant  $\hat{\Phi}$  that appears in eq. (40). On the other hand this quantity has a different physical interpretation. Indeed, the conformation matrix  $\tilde{\mathbf{Q}}^c$  describes transitions between metastable states, while the matrix  $\tilde{\mathbf{Q}}$  describes transitions between neighboring states.

### 4.1 Two dimensional system

We consider a two-dimensional convection-diffusion process of the stochastic differential equation:

$$\begin{cases} dx_t = -\nabla_x V(x_t, y_t) + \sigma dB_t^x \\ dy_t = -\nabla_y V(x_t, y_t) + \sigma dB_t^y \end{cases} \quad (69)$$

where  $B_t^i$  denotes a standard Brownian motion in the direction  $i = x, y$ ,  $\sigma = 1, 1.5, 2, 2.5$  is the volatility and  $V(x_t, y_t)$  is a two-dimensional potential energy surface given by the function

$$V(x, y) = 4 \left( x^3 - \frac{3}{2}x \right)^2 - x^3 + x + 2 + 2y^2 \quad (70)$$

This function describes a two-dimensional triple-well potential along the direction  $x$  as reported in fig. (6) and guarantees the presence of three metastable states. The system has been solved by using the Euler-Maruyama scheme:

$$\begin{cases} x_{n+1} = x_n - \nabla_x V(x_n, y_n) \Delta t + \sigma \eta^x \sqrt{\Delta t} \\ y_{n+1} = y_n - \nabla_y V(x_n, y_n) \Delta t + \sigma \eta^y \sqrt{\Delta t} \end{cases} \quad (71)$$

where  $\Delta t = 0.001$  is the integration time-step and  $\eta^i$  are i.i.d random variables drawn from a standard Gaussian distribution. We have produced trajectories of  $4 \times 10^7$  time-steps.

The Voronoi tessellation has been constructed such that the cells are of approximately the same size. We chose a random initial point  $c_0$  of the trajectory as first center of a cell, and fixed a minimum distance  $r$ . To select the next center  $c_1$ , we have picked the point of the trajectory nearest to  $c_0$  outside the radius  $r$ . Then we iterated this procedure to search the next centers. Fig. 5 describes the algorithm. Note that one can speed up the search, by removing all points from the list that are within a radius  $r$  from the already found centers.

This algorithm guarantees the cells are homogeneously distributed and have approximately the same area. The final number of cells depends on the temperature of the system. A simulation at high temperature samples a bigger state space than a simulation at low temperature, thus also the number of Voronoi cells, whose volume depend on the variable  $r$ , increases. We have used  $r = 0.1, 0.15, 0.2$  that yield between 400 and 2700 cells depending on the volatility (temperature) of the system.

The set of the points  $c_i$  is used to construct the Voronoi tessellation. Actually this operation could be avoided, because we need to know only the adjacency matrix that identifies the neighboring cells. Thinking the Voronoi diagram in term of convex polyhedra[6, 14] permits to write a Linear Program to estimate the adjacency matrix, improving the efficiency compared to usual algorithms.

The rate matrix  $\tilde{\mathbf{Q}}$  has been constructed according to the equation (66) with  $\beta = 2/\sigma^2$ . The MSM has been constructed on the same tessellation choosing a lag time range from 100 to 1000 time steps. The PCCA+ analysis has been realized by using  $n_c = 3$  (the number of metastable states) as input parameter.

To study the dependence of the flux on the potential energy, we have perturbed the potential energy function (70) by adding the function

$$U(\kappa, x) = \kappa x \quad (72)$$

where  $\kappa$  is a parameter that tunes the strength of the perturbation.

## 4.2 Alanine dipeptide

We performed all-atom MD simulations of acetyl-alanine-methylamide (Ac-A-NHMe, alanine dipeptide) in explicit water. All simulations were carried out with the GROMACS

5.0.2 simulation package[20] with the force field AMBER ff-99SB-ildn [15] and the TIP3P water model [10]. We have performed simulations in a NVT ensemble, at temperature of 900 K and at temperture of 300 K. The length of each simulation was 600 ns and we printed out the positions every `nstxout=500` time steps, corresponding to 1 ps. A velocity rescale thermostat has been applied to control the temperature and a leapfrog integrator has been used to integrate the equation of the motion. To discretize the state space with a Voronoi tessellation, we used the same algorithm described for the two-dimensional system. The same discretization has been used also to construct the MSM.

## 5 Results and discussion

### 5.1 Two dimensional system

As first application, we consider a two dimensional diffusion process. This example is important to highlight the main properties of the rate matrix estimated by square root approximation. The potential energy function (eq. (70), fig. 6) has three minima respectively at (-1.12,0), (0.05, 0) and (1.29,0), separated by three barriers, whose highest points are approximatively at (-0.83, 0) and (0.61,0).

Fig. 7 shows three different trajectories generated respectively with volatility  $\sigma = 1.0$ ,  $\sigma = 1.5$  and  $\sigma = 2.0$  together with the Voronoi discretization of the space, keeping constant the parameter  $r$ , i.e. the minimum distance between the centers of neighboring cells. Note that this parameter is also related to the number of cells and to their volume. In particular, if the state space is uniformly sampled, the minimum distance between neighboring centers turns out to be approximately the average distance between neighboring centers. This guarantees also that the cells have on average the same volume. By increasing the volatility, that represents the temperature of the environment, the trajectory covers a larger area, thus the Voronoi tessellation results to have more cells, distributed more homogeneously and with almost the same volume. As we will see, this fact has big impact on the construction of the rate matrix.

We start discussing the results of a simulation realized with volatility  $\sigma = 1.5$ , after having discretized the space in 1258 Voronoi cells, using a minimum distance between centers of neighbor cells  $r = 0.1$  and taking into account periodic boundary condition. Fig. 8 shows the Boltzmann weights of the Voronoi cells, estimated as average over all the points falling in each cell:

$$\pi_i = \frac{1}{\text{nsteps}} \sum_{n=1}^{\text{nsteps}} \mathbf{1}_{\Omega_i}(x_n, y_n) \exp(-\beta V(x_n, y_n)) \quad \forall \text{cell } \Omega_i \quad (73)$$

where  $\mathbf{1}_{\Omega_i}(x_n, y_n)$  is the indicator function and `nsteps` is the length of the trajectory. The picture of the Boltzmann weights represents the distribution of the most visited states and it is specular to the potential energy function. Thus, we observe high values in correspondence of the potential minima, and low values in correspondence of the barriers. The Boltzmann weights have been used to estimate the matrix (66)  $\tilde{\mathbf{Q}}$  that, by definition is proportional to the rate matrix  $\mathbf{Q}'''$  in eq. (42), although we still do not know the value of the proportionality constant  $\hat{\Phi}$ .

This matrix, even if it does not contain the correct rates, can be used for a qualitative analysis of the dynamics of the system. Fig. 9 shows the first six left eigenvectors of

the matrix. The first eigenvector has only positive entries and is proportional to the Boltzmann weights in fig. 8. The second eigenvector contains also negative values and represents the slowest transition of the dynamics that happens between the third minimum (the deepest minimum) and one of the two left minima. The third eigenvector represents the second slowest transition between the central and external minima. The next eigenvectors describe the faster processes that can happen even between states belonging to the same metastable state.

The eigenvalues of the matrix contain information about the implied timescales, i.e. the average times at which the transitions occur. The first eigenvalue is zero, according to the Perron-Frobenius theorem, while the others are negative. It is known[17] that the implied timescales of a Markovian process, described by a transition probability matrix  $\mathbf{T}(\tau)$  are obtained as

$$t_i = -\frac{\tau}{\log \lambda_i(\tau)} \quad (74)$$

where  $\lambda_i(\tau)$  is the  $i$ th eigenvalue of the matrix  $\mathbf{T}(\tau)$ . Thus, exploiting the semigroup property (63), a simple calculation shows that the implied timescales can be obtained as

$$t_i = -\frac{1}{\hat{\Phi}\tilde{\theta}_i} \quad (75)$$

where  $\tilde{\theta}_i$  are the eigenvalues of the rate matrix  $\tilde{\mathbf{Q}}$ . The advantage of the relation (75) respect to (74) is that one does not need anymore to check the convergence of the implied timescales as in Markov State Model construction [17], because eq. (75) does not depend on the lag-time  $\tau$ . On the other hand we do not know the value of  $\hat{\Phi}$  yet and thus we can only relatively compare the transitions one to another. For example, the first slow transition happens at a timescale of  $t_1 = 70.53\hat{\Phi}^{-1}$  time steps, while the second transition at  $t_2 = 19.14\hat{\Phi}^{-1}$  time steps, then the second transition is 3.68 times faster than the first one. Table 1 contains the first five implied timescales, while the ratios between the implied timescales are in table 2.

To estimate the value of the flux, we have constructed the transition probability matrix by Markov State Model, then we have reduced both the matrices  $\tilde{\mathbf{Q}}$  and  $\mathbf{T}(\tau)$  by PCCA+, to the conformation matrices  $\tilde{\mathbf{Q}}^c$  and  $\mathbf{T}^c(\tau)$ . Afterwards, we have estimated the constant  $\hat{\Phi}$  by studying the eigenvalues, according to equation (68).

Because the system has three metastable states, the matrices  $\tilde{\mathbf{Q}}^c$  and  $\mathbf{T}^c(\tau)$  are  $3 \times 3$ , thus the constant  $\hat{\Phi}$  can be estimated comparing the eigenvalues  $\tilde{\theta}_i$  and  $\lambda_i(\tau)$  with  $i = 2, 3$ . Tables 3, 4, 5 and 6 contain the values of  $\hat{\Phi}$  for several simulations with different initial conditions. In each table,  $\hat{\Phi}_2$  denotes the flux estimated comparing the second eigenvalues,  $\hat{\Phi}_3$  denotes the flux estimated comparing the third eigenvalues,  $\bar{\Phi}$  is the average flux between  $\hat{\Phi}_2$  and  $\hat{\Phi}_3$ , "std" is the standard deviation and "rel. err." is the relative error.

Table 3 reports the flux as function of the volatility  $\sigma$ , i.e. the temperature of the environment. We observe that the average flux increases linearly. Indeed at high temperature, the dynamics is faster and a single cell is visited for a short amount of time. Furthermore, the trajectory covers also those zones difficult to sample due to a high gradient of the potential energy function. As consequence, the number of Voronoi cells grows up (keeping constant the parameter  $r$ ) and we observe a reduction of the relative error.

In table 4 we observe how the flux depends on the choice of the parameter  $r$ , i.e. on the minimum distance between the centers of neighboring cells. By reducing  $r$  (and

consequently the volume of the cells), the average flux increases. Indeed by reducing the size of the cells, a cell is visited for a minor amount of time and the number of transitions through the intersecting surfaces grows up. We observe also an important reduction of the relative error that denotes the convergence of the rate matrix.

Finally, we have studied the dependence of the flux on the external perturbation of the potential energy function described in eq. (72) with  $\kappa = 0, 0.5, 1$ . The effect of such perturbation is to tilt the potential energy along the axis  $x$ , and consequently it redistributes the Boltzmann weights. The experiment has been repeated for  $\sigma = 1.5$  and  $\sigma = 2.0$  and the results are collected respectively in table 5 and 6. In both the cases the perturbation does not affect the average flux. This result confirms the initial assumption in section 2.5, that the flux does not depend on the potential energy function.

## 5.2 Alanine dipeptide

We have repeated a similar analysis for alanine dipeptide (Ac-A-NHMe) in explicit water. The main difference with respect to the 2d system is that the simulation has been carried out on all the degrees of freedom (all atomic simulation), but we have constructed the rate matrix (and the MSM) only on two relevant coordinates, the backbone dihedral angles  $\phi$  and  $\psi$ , that well capture the main dynamical properties of the system. This approximation requires to fix the proposed approach to obtain the rate matrix in a low dimension space.

First of all, to get the correct value of the Boltzmann weights one should use the potential energy function  $V(\mathbf{r})$  (force field) in full dimension. However, this would return an unreadable result due to the high dimensionality of the system, thus we have replaced the true Boltzmann weights in full dimension, with a normalized histogram on the dihedral angles  $\phi$  and  $\psi$ . This assumption is correct if the relevant coordinates well represent the stationary distribution of the system. In such case, we have

$$\pi_i = \mathbf{1}_{\Omega_i}(\mathbf{r}) \frac{\exp(-\beta V(\mathbf{r}))}{Z} \propto h_i = \frac{\sum_{n=1}^{\text{nsteps}} \mathbf{1}_{\Omega_i}(\phi_n, \psi_n)}{\text{nsteps}} \quad \forall \text{cell } \Omega_i \quad (76)$$

where  $\pi_i$  is the true Boltzmann weight of the cell  $\Omega_i$ ,  $Z$  is the unknown partition function,  $\{\phi_n, \psi_n\}$  is the state of the trajectory at time step  $n$  and  $\text{nsteps}$  is the length of the trajectory. The histogram can be used also to estimate the free energy profile  $F(\phi, \psi)$ , defined up to a negligible constant and to rescale the Boltzmann weights according to a new temperature.

$$\begin{aligned} h_i &= \frac{\exp(-\beta F_i)}{N} \quad \forall \text{cell } \Omega_i \\ F_i &= -\frac{1}{\beta} \log h_i + \text{const.} \end{aligned} \quad (77)$$

where  $F_i$  is the free energy of the cell  $\Omega_i$  in the space  $\{\phi_n, \psi_n\}$  and  $N$  is a normalization constant. The new Boltzmann weights, for a different temperature  $\beta' = 1/(k_B T')$  would be

$$h'_i = \frac{\exp(-\beta' F_i)}{N} \quad \forall \text{cell } \Omega_i \quad (78)$$

Let's now discuss the results. Fig. 10 shows respectively two trajectories generated at temperature 300 K (A and B) and 900 K (C), projected on the backbone dihedral angles

$\phi$  and  $\psi$ . At temperature 300 K, the trajectory does not cover the full space and the Voronoi tessellation presents cells with different size. Reducing the parameter  $r$ , does not improve the discretization. Even though the number of cells grows (and the size reduces), there are still cells with very large volume. The only way to improve the discretization, i.e. to have small cells with almost the same volume, is to increase the temperature to promote a better sampling of the state space. Indeed, the trajectory at temperature 900 K covers a larger subspace of the state space, that turns out in a better Voronoi discretization of the space.

Fig. 11 shows the Boltzmann weights at temperature 900 K, that correspond to the typical equilibrium distribution in the Ramachandran plane. Fig. 12 shows the dominant left eigenvectors of the matrix  $\tilde{Q}_{ij} = \sqrt{h_j/h_i}$ . The entries of the first eigenvector are only positive and proportional to the Boltzmann weights  $h_i$ . The second eigenvector represents a torsion around the  $\phi$  angle and corresponds to a kinetic exchange between the  $L_\alpha$ -minimum ( $\phi > 0$ ) and the  $\alpha$ -helix and  $\beta$ -sheet minima ( $\phi < 0$ ). The associated transition, according to eq. (75), is  $146\hat{\Phi}^{-1}$  ps. The third eigenvector represents a transition  $\beta$ -sheet  $\longleftrightarrow \alpha$ -helical conformation, i.e. a torsion around  $\psi$ , and is associated to a timescale of  $40\hat{\Phi}^{-1}$  ps. The eigenvectors match very well with the eigenvectors that one would obtain from a MSM[21], however we do not know the value of the constant  $\hat{\Phi}$ , thus the timescales need to be corrected.

As in the two-dimensional case, we have constructed a MSM, then we have produced the conformation matrices assuming the existence of three metastable states ( $\beta$ -sheet,  $L\alpha$ -helix and  $R\alpha$ -helix) and we have compared the eigenvalues.

Table 7 collects the results for different experiments. In the first two rows of the table, we used the same trajectory at high temperature, to construct the rate matrix and the MSM. We observe that increasing the number of cells (i.e. by reducing their volume), the average flux grows up and the precision improves, thus we have that  $\hat{\Phi}_2 \simeq \hat{\Phi}_3$ . This result confirms what we already discussed for the two-dimensional diffusion process.

The third, forth and fifth rows contain the results for a trajectory generated at 300 K. In this case we observe that by increasing the number of cells causes an increase of the flux, however, the relative error is very high in all the three cases. As we can see in fig. 10.A, alanine dipeptide presents 3 high barriers (approximately at  $\phi \approx 0$ ,  $\phi \approx 2/3\pi$  and  $\psi \approx -4/5\pi$ ) that are not well sampled at low temperature. In particular we observe very large cells covering the zones of the Ramachandran plane (corresponding to high barriers) not visited by the trajectory. Reducing the parameter  $r$  (and then the size of the cells), does not improve the result (the relative error is still very high), because that parameter does not affect the size of those cells in not visited regions. This results in a violation of the initial condition and in a bad rate matrix.

The last two rows of the table show the results obtained comparing the eigenvalues of a rate matrix built on the trajectory at temperature  $T = 900$  K, with eigenvalues of the transition probability matrix, built on the trajectory at temperature  $T = 300$  K. We remark that we rescaled the rate matrix to  $T = 300$  K, according to equation (78). Also in this case, by increasing the number of cells, we can observe a significant increase of the flux. The relative error is approximately halved with respect to the third, forth and fifth row. We deduce that this improvement is due to the better rate matrix, constructed with a trajectory produced at  $T = 900$  K. On the other hand, the error is still high, because the transition probability matrix has been constructed between states not well sampled (i.e. from the trajectory at  $T = 300$  K).

We now propose a scheme to improve the result at low temperature.

1. Run a Metadynamics simulation at  $T = 300$  K, perturbing the relevant coordinates  $\phi, \psi$ . This will produce a free energy profile  $F(\phi, \psi)$  as sum of Gaussian functions.
2. Construct the rate matrix with the found free energy profile. Metadynamics guarantees a trajectory that covers the full Ramachandran plane, thus the Voronoi discretization will not be affected by not visited regions and the resulting rate matrix will improve the result.
3. Construct a MSM with the metadynamics trajectory, then reweight the MSM by Girsanov theorem. This procedure guarantees a better MSM with respect to the standard method, because the transition probabilities between rarely visited states are now available.
4. Obtain the flux from the eigenvalues of the rate matrix and the reweighted MSM.

## 6 Conclusion

The paper contributes to the classical molecular simulation community in three ways. It provides an easy way to estimate the rates between metastable molecular conformations. It shows that this type of discretization converges to a Fokker-Planck-operator. Finally, it shows that there is an easy mathematical relation between the discretized generator of the molecular process and the potential energy landscape.

More detailed: Since many years the concept of transfer operators, well known in thermodynamics and quantum mechanics, has been established inside the classical molecular simulation community. New methods, such as Markov State Models have been developed, to reduce complexity and study conformational transition networks of molecular systems. The concept of transfer operators is strongly connected to the concept of generators. Numerical methods which apply for the transfer operator also can be used for the generator, which is simply the time-derivative of the transfer operator. The spatial discretization of the transfer operator is a transition probability matrix of a Markov chain. The spatial discretization of the generator is a rate matrix, which is in general hard to extract from time-discretized simulation data. Our method simply uses the Boltzmann distribution of states for discretizing the generator. The first underlying assumption is that we can define a continuity equation for the time-derivative of the transfer operator. Then, exploiting the Gauss theorem, we write the rate between two neighbor states as a surface integral of the flux, weighted by the Boltzmann density of the intersecting surface. The second assumption is a constant flux, i.e, the flux does not depend on the potential energy but on the discretization of the space. Instead of computing the Boltzmann weight of the intersecting surface of the Voronoi cells, here, it is estimated as geometric average of the Boltzmann weight of the cells. This we denoted as square root approximation. The quality of the square root approximation depends on the discretization of the state space. We propose to use a Voronoi tessellation. For this we proved that the rate matrix converges to the generator of the Smoluchowski equation[7]. In fact, the numerical and molecular examples show that the corresponding rate matrices converge to a Fokker-Planck operator as the number of cells increases.

The constant flux trough the surfaces of the Voronoi cells cannot be computed directly. By comparing the eigenvalues of the rate matrix with the corresponding eigenvalues of the transition probability matrix estimated by MSM a good estimate for the flux can be provided. This especially accounts for an increasing number of cells.

According to the theory, a fine discretization improves the final result and reduces the relative error of the flux computation. Furthermore, we have shown that an external perturbation of the potential energy function does not affect the result, according to the second assumption. In both the systems studied, we have observed that the quality of the rate matrix depend on the quality of the sampling. This is in accordance with theoretical studies[7]. High temperature simulations results are more suitable to estimate the rate matrix because of the better mixing properties. A high temperature (higher entropy) also results in more homogeneous Voronoi cells.

## A Notation

<code>nsteps</code>	length of a trajectory
$\Gamma \subset \mathbb{R}^{6N}$	phase space of the system
$q(t), p(t)$	position and momentum at time $t$
$\mathbf{x}(t) = \{q(t), p(t)\}$	state of the system at time $t$
$\mathcal{H}(\mathbf{x})$	Hamiltonian function
$V(q)$	potential energy function
$B$	brownian motion
$\rho_t(\mathbf{x})$	probability density function at time $t$
$p(\mathbf{x}, \mathbf{y}, t)$	transition kernel
$\mu(\mathbf{x})$	invariant measure respect to $p(\mathbf{x}, \mathbf{y}, t)$
$\pi(\mathbf{x})$	Boltzmann density function
$\mathcal{P}(t)$	propagator
$\mathcal{T}_{fwd}(t)$	forward transfer operator
$\mathcal{T}_{bwd}(t)$	backward transfer operator
$\mathcal{Q}_{fwd}$	forward infinitesimal generator
$\mathcal{Q}_{bwd}$	backward infinitesimal generator
$\mathbf{T}(t)$	Galerkin discretization of $\mathcal{T}_{bwd/fwd}(t)$ (transition probability matrix)
$\mathbf{C}(\tau)$	counting matrix to construct MSM
$\mathbf{Q}$	Galerkin discretization of $\mathcal{Q}_{bwd}$ (true rate matrix)
$\mathbf{Q}'$	rate matrix assuming constant flux
$\mathbf{Q}''$	rate matrix assuming that the flux is concentrated in a point of the intersecting surface
$\tilde{\mathbf{Q}}''$	rate matrix after square root approximation
$\tilde{\mathbf{Q}}$	rate matrix after square root approximation, neglecting the constant flux
$\mathbf{T}^c(\tau)$	Galerkin discretization of $\mathcal{T}_{bwd}(t)$ in the space of the conformations
$\mathbf{Q}^c$	Galerkin discretization of $\mathcal{Q}_{bwd}$ in the space of the conformations
$\Lambda$	eigenvalues of $\mathbf{T}(\tau)$
$\Theta$	eigenvalues of $\mathbf{Q}$
$f$	eigenfunctions of $\mathcal{T}_{bwd}(t)$ and $\mathcal{Q}_{bwd}$
$X$	eigenvectors of $\mathbf{T}(t)$ and $\mathbf{Q}$
$\mathbf{A}$	transformation matrix (PCCA+)
$\chi_i$	membership functions
$\tau$	lag-time
$\Omega$	Voronoi tessellation of $\Gamma$
$\Omega_i$	Voronoi cell
$\partial\Omega_i \partial\Omega_j$	common area of the Voronoi cells $i$ and $j$
$dS$	infinitesimal piece of the surface $S$
$\mathbf{j}$	flux density
$\Phi(\mathbf{x})$	flux of the configuration $\mathbf{x}$ trough the surface of a Voronoi cell
$\hat{\Phi}$	constant flux. Assumption to construct the square root approximation

$\hat{\Phi}_{2,3}$	constant flux from the second and third eigenvalues
$\bar{\Phi}$	average constant flux between $\hat{\Phi}_2$ and $\hat{\Phi}_3$
$T$	temperature of the system
$\sigma$	volatility of the diffusion process

## References

- [1] Everything you wanted to know about markov state models but were afraid to ask. *Methods*, 52(1):99–105, 2010.
- [2] Roberto Alicandro, Marco Cicalese, and Antoine Gloria. Integral representation results for energies defined on stochastic lattices and application to nonlinear elasticity. *Archive for Rational Mechanics and Analysis*, 200(3):881–943, 2011.
- [3] M. Biskup. Recent progress on the random conductance model. *Probab. Surv.*, 8:294–373, 2011.
- [4] J.-P. Bouchaud and A. Georges. Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications. *Phys. Rep.*, 195(4-5):127–293, 1990.
- [5] Peter Deuflhard and Weber Marcus. Robust perron cluster analysis in conformation dynamics. *Linear Algebra and its Applications*, 398:161–184, 2004.
- [6] Komei Fukuda. Frequently asked questions in polyhedral computation, 2004.
- [7] M. Heida. On convergences of the squareroot approximation scheme to the fokker-planck operator. *WIAS Preprint Nr.*
- [8] Wilhelm Huiszinga. *Metastability of Markovian systems*. PhD dissertation, Freien Universität Berlin, Fachbereich Mathematik und Informatik, May 2001.
- [9] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik. *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, 1994. Translated from the Russian by G. A. Yosifian [G. A. Iosif'yan].
- [10] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M.L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926, 1983.
- [11] Susanna Kube and Marcus Weber. Conformation kinetics as a reduced model for transition pathways. *ZIB Report*, 43(September):99–36, 2005.
- [12] Susanna Kube and Marcus Weber. Coarse grained molecular kinetics. *ZIB Report*, 35, 2006.
- [13] Susanna Kube and Marcus Weber. Preserving the markov property of reduced reversible markov chains. *AIP Conference Proceedings*, 593(1048), 2008.
- [14] Han Cheng Lie, Konstantin Fackeldey, and Marcus Weber. A square root approximation of transition rates for a markov state model. *SIAM. J. Matrix Anal. Appl.*, 34:738–756, 2013.
- [15] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, and E. Shaw, D. Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins*, 78:1950–1958, 2010.
- [16] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer Verlag, Berlin, 6th edition, 2003.

- [17] Jan-Hendrik Prinz, Hao Wu, Marco Sarich, Bettina Keller, Martin Senne, Martin Held, John D Chodera, Christof Schütte, and Frank Noé. Markov models of molecular kinetics: generation and validation. *J. Chem. Phys.*, 134:174105, 2011.
- [18] Axel Schild. *Electron Fluxes During Chemical Processes in the Electronic Ground State*. PhD thesis, Freie Universität Berlin, Berlin, 2013. [http://www.diss.fu-berlin.de/diss/receive/FUDISS\\_thesis\\_00000095279](http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_00000095279).
- [19] Christof Schütte, Wilhelm Huisinga, and Peter Deuflhard. Transfer operator approach to conformational dynamics in biomolecular systems. *ZIB Report*, 36(October):SC 99–36, 1999.
- [20] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, E. Mark, A., and J. Berendsen, H. Gromacs: fast, flexible, and free. *J. Comput. Chem.*, 26(16):1701–1719, 2005.
- [21] F Vitalini, A S J S Mey, F Noé, B G Keller, F Vitalini, A S J S Mey, F Noé, and B G Keller. Dynamic properties of force fields. *J. Chem. Phys.*, 142:084101, 2015.
- [22] Marcus Weber. An efficient analysis of rare events in canonical ensemble dynamics. *ZIB Report*, 611, 2009.
- [23] Marcus Weber. *A Subspace Approach to Molecular Markov State Models via a New Infinitesimal Generator*. Habilitation thesis, Freien Universität Berlin, Zuse Institute Berlin, 2011.

## Tables

$i$	$\theta_i$	I.T.S. (time steps)
0	0	-
1	-0.0142	$70.5319 \hat{\Phi}^{-1}$
2	-0.0522	$19.1482 \hat{\Phi}^{-1}$
3	-0.0738	$13.5452 \hat{\Phi}^{-1}$
4	-0.0871	$11.4772 \hat{\Phi}^{-1}$
5	-0.1252	$7.9900 \hat{\Phi}^{-1}$

Table 1: First five eigenvalues and implied timescales as function of the flux. The volatility was set equal to 1.5 and the minimum distance between the centers of neighbor cells was set equal to  $r = 0.1$ .

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
$i = 1$	1.00	3.68	5.20	6.14	8.82
$i = 2$	0.27	1.00	1.41	1.66	2.39
$i = 3$	0.19	0.70	1.00	1.18	1.69
$i = 4$	0.16	0.59	0.84	1.00	1.43
$i = 5$	0.11	0.41	0.58	0.69	1.00

Table 2: Ratio  $ITS_i/ITS_j$  between the first five implied timescales of table 1. The volatility was set equal to 1.5 and the minimum distance between the centers of neighbor cells was set equal to  $r = 0.1$ .

$\sigma$	ncells	$\hat{\Phi}_2$	$\hat{\Phi}_3$	$\bar{\Phi}$	std	rel. err.
1.0	740	0.0080	0.0161	0.0121	0.0057	47,11%
1.5	1258	0.0461	0.0523	0.0492	0.0044	8.94%
2.0	1725	0.0913	0.0931	0.0922	0.0013	1.41%
2.5	2205	0.1427	0.1372	0.1400	0.0039	2.79%

Table 3: Variation of the flux as function of the volatility  $\sigma$ . The minimum distance between the centers of neighbor cells was set equal to  $r = 0.1$ .

$r$	ncells	$\hat{\Phi}_2$	$\hat{\Phi}_3$	$\bar{\Phi}$	std	rel. err.
0.20	456	0.0207	0.0248	0.0227	0.0029	12.87%
0.15	784	0.0396	0.0425	0.0410	0.0021	5.12%
0.10	1725	0.0913	0.0931	0.0922	0.0013	1.41%

Table 4: Variation of the flux as function of the minimum distance between the centers of neighbor cells. The volatility was set equal to 2.0.

$\kappa$	n <b>cells</b>	$\hat{\Phi}_2$	$\hat{\Phi}_3$	$\bar{\Phi}$	std	rel. err.
0.0	1258	0.0461	0.0523	0.0492	0.0044	8.94%
0.5	1235	0.0459	0.0502	0.0480	0.0030	6.25%
1.0	1225	0.0449	0.0526	0.0487	0.0055	11.29%

Table 5: Variation of the flux as function of an external perturbation. The volatility was set equal to 1.5 and  $r = 0.1$ .

$\kappa$	n <b>cells</b>	$\hat{\Phi}_2$	$\hat{\Phi}_3$	$\bar{\Phi}$	std	rel. err.
0.0	1725	0.0913	0.0931	0.0922	0.0013	1.41%
0.5	1722	0.0927	0.0922	0.0924	0.0003	0.36%
1.0	1720	0.0904	0.0926	0.0915	0.0015	1.64%

Table 6: Variation of the flux as function of an external perturbation. The volatility was set equal to 2.0 and  $r = 0.1$ .

Temp. (K)	$r$	n <b>cells</b>	$\hat{\Phi}_2$	$\hat{\Phi}_3$	$\bar{\Phi}$	std	rel. err.
900	0.20	740	11.4390	12.9166	12.1778	1.0448	8.58%
900	0.17	1005	17.0914	16.5862	16.8388	0.3572	2.12%
300	0.17	566	0.1238	2.5672	1.3455	1.7277	128.41%
300	0.14	792	0.1645	3.0356	1.6000	2.0301	126.88%
300	0.10	1423	0.3090	6.4278	3.3684	4.3266	128.45%
$\theta_i(900) \sim \lambda_i(300)$	0.17	1017	6.6788	2.5062	4.5925	2.9505	64.24%
$\theta_i(900) \sim \lambda_i(300)$	0.10	2660	15.6359	6.1802	10.9081	6.6862	61.30%

Table 7: Alanine dipeptide.

## Figures

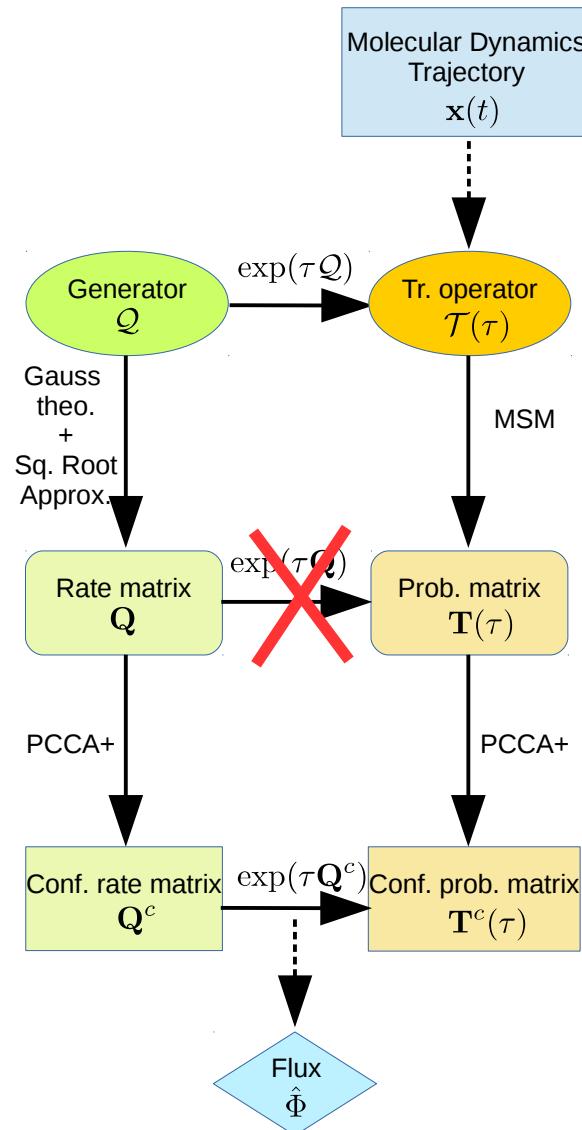


Figure 1: Scheme summarizing the theory of the transfer operator, the generator and the associated transition matrices.

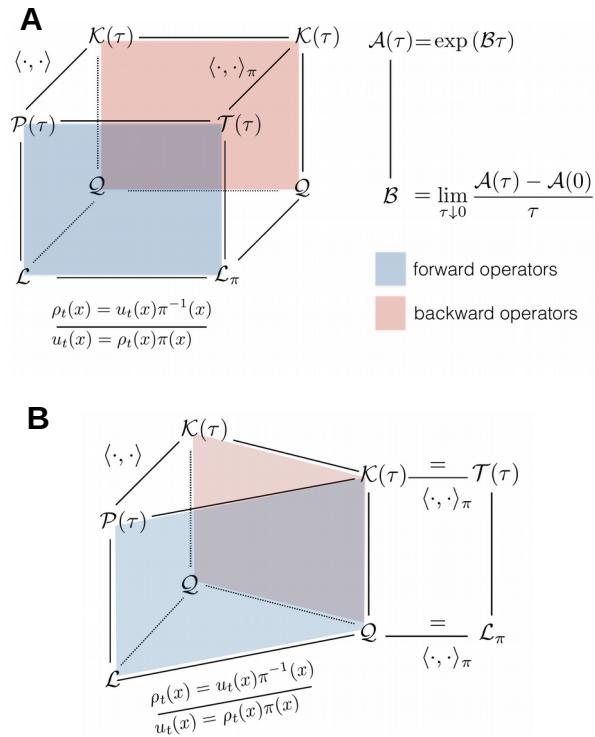


Figure 2: Scheme summarizing the theory of the transfer operator, the generator and the associated transition matrices.

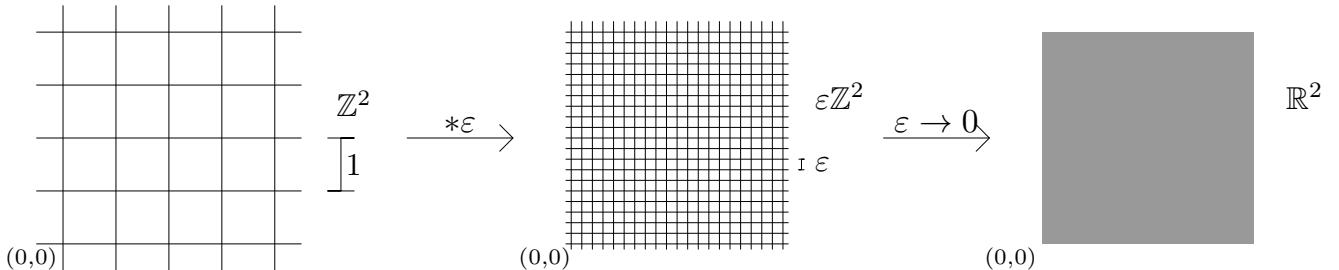


Figure 3: The role of the parameter in the case  $\varepsilon\mathbb{Z}^2$ . As  $\varepsilon \rightarrow 0$ , the grid becomes finer and finer and approximates the whole of  $\mathbb{R}^2$ .

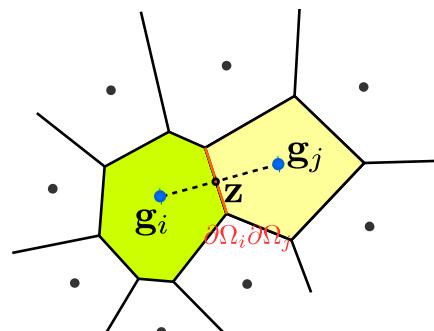


Figure 4: Square root approximation.

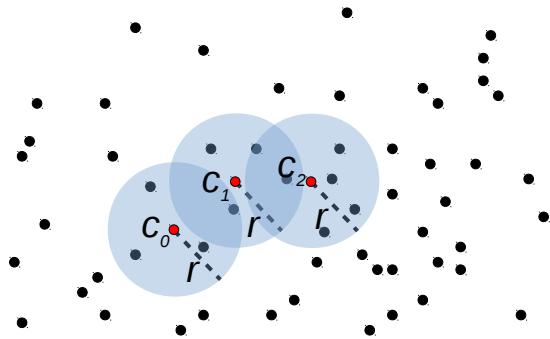


Figure 5: Selection of the states to make the Voronoi discretization. The black points represent the states of a trajectory, to select those are becoming centers of the Voronoi cells, we have set a minimum distance  $r$ . The first center  $c_0$  is randomly chosen, the next center  $c_1$  is the nearest out of a radius  $r$ . The next centers are found iteratively.

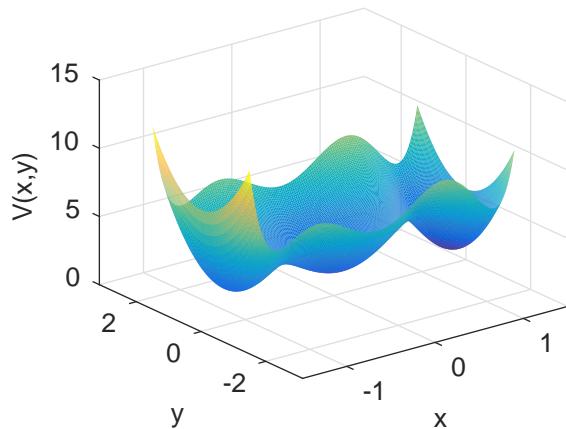


Figure 6: Diffusion process 2d. Potential energy function  $V(x, y)$ .

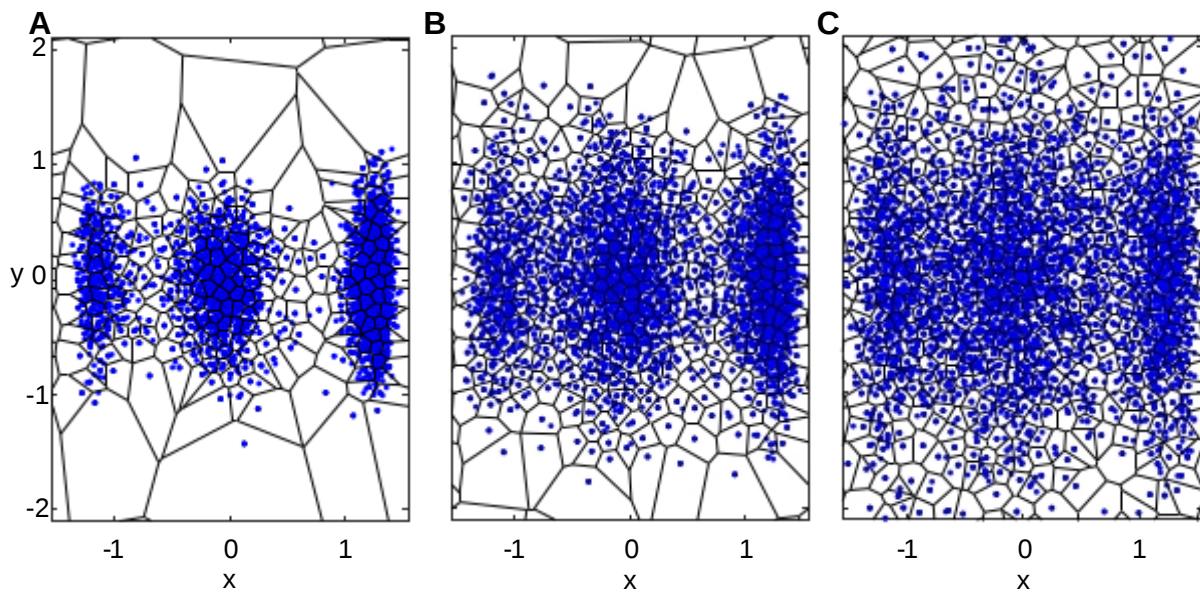


Figure 7: Diffusion process 2d. Trajectories  $(x_n, y_n)$  generated with  $\sigma = 1.0$  (A),  $\sigma = 1.5$  (B) and  $\sigma = 2.0$  (C). The Voronoi tessellation are made respectively by 251, 452 and 605 cells. To realize the pictures, we picked  $4 \times 10^4$  points, every 1000 timesteps, from long trajectories of  $4 \times 10^7$  points.

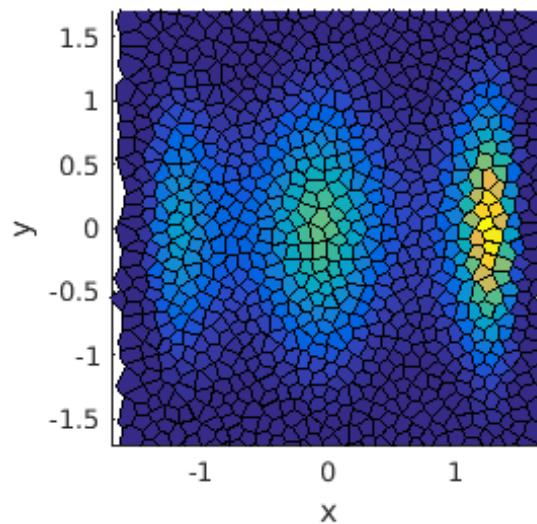


Figure 8: Diffusion process 2d. Boltzmann density of the Voronoi cells (`ncells=1258`). The gradient color, from blue to yellow denotes the density of the cells. The trajectory used to construct the rate matrix has been generated with  $\sigma = 1.5$  and  $r = 0.1$ .

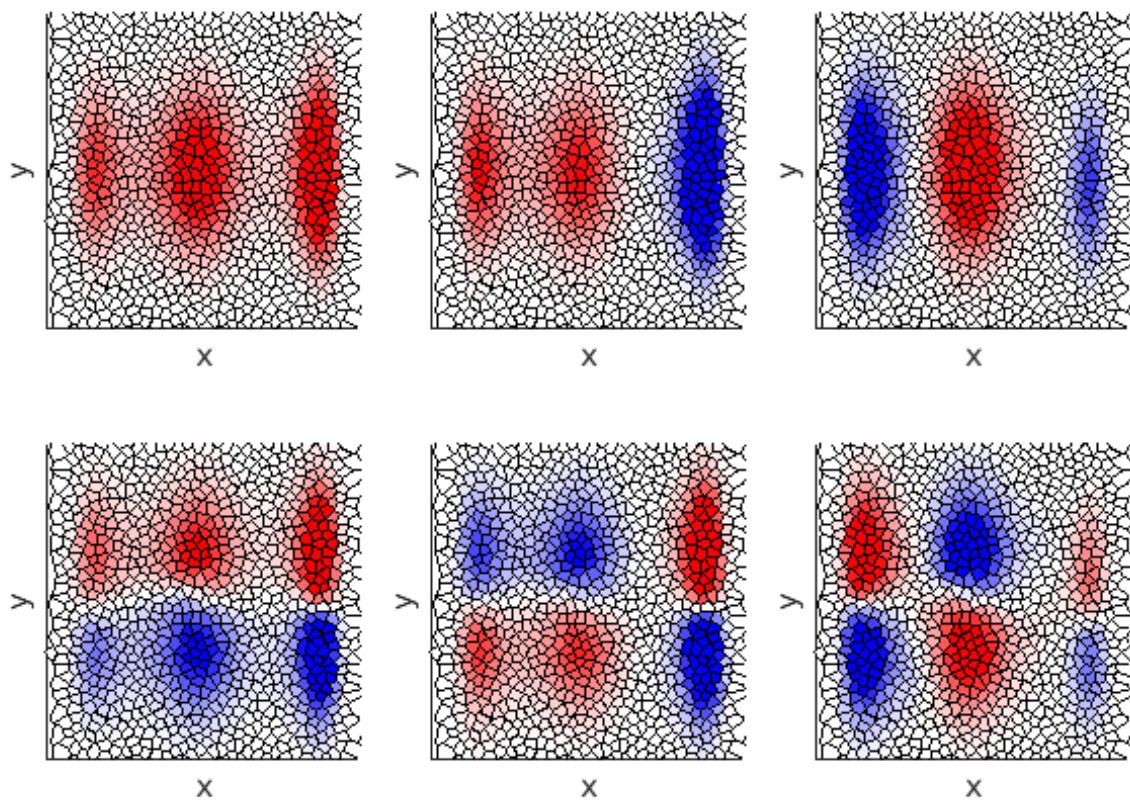


Figure 9: Diffusion process 2d. Eigenvectors of the rate matrix  $\tilde{\mathbf{Q}}$ . The red color denotes positive entries, the blue color denotes negative entries. The trajectory used to construct the rate matrix has been generated with  $\sigma = 1.5$  and  $r = 0.1$ .

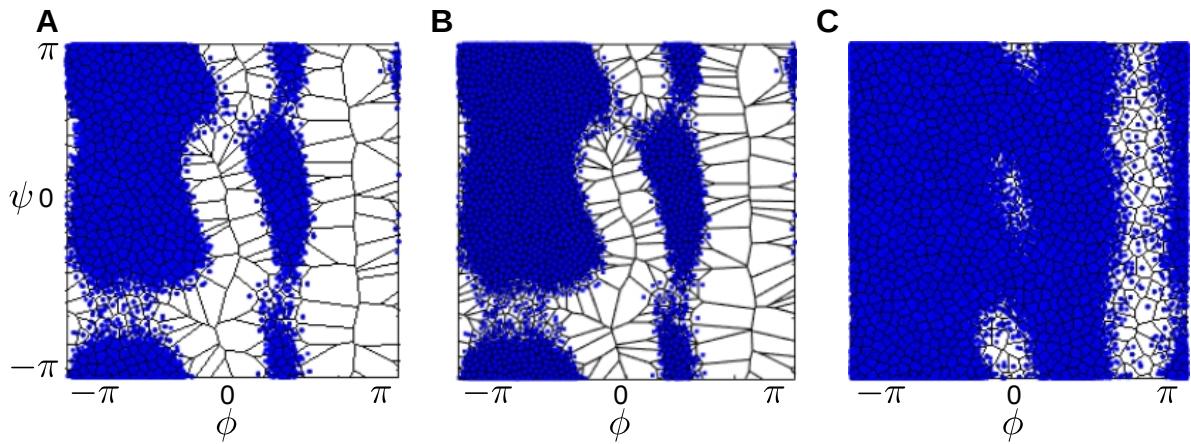


Figure 10: Alanine dipeptide. Trajectories  $(\phi_n, \psi_n)$  generated at temperature 300 K (A) and (B), at 900 K (C). The Voronoi tessellation are made respectively by 566 (A), 1433 (B) and 1005 cells, using  $r = 0.17$  (A and C),  $r = 0.1$  (B).

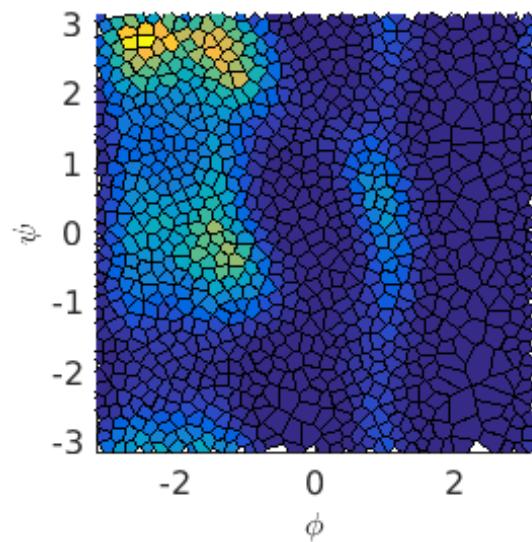


Figure 11: Alanine dipeptide. Boltzmann density of the Voronoi cells ( $n_{\text{cells}}=1005$ ). The gradient color, from blue to yellow denotes the density of the cells. The trajectory used to construct the rate matrix has been generated at 900 K and  $r = 0.17$ .

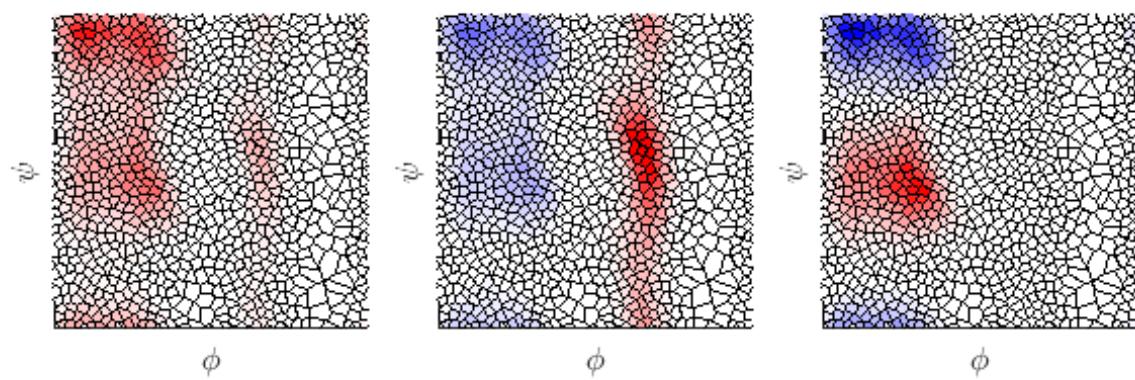


Figure 12: Alanine dipeptide. Eigenvectors of the rate matrix  $\tilde{\mathbf{Q}}$ . The red color denotes positive entries, the blue color denotes negative entries. The trajectory used to construct the rate matrix has been generated at 900 K and  $r = 0.17$ .