

MARCUS WEBER

**An efficient analysis of rare events in canonical
ensemble dynamics**

An efficient analysis of rare events in canonical ensemble dynamics

Marcus Weber

Zuse-Institute Berlin, Takustraße 7, D-14195 Berlin, Germany

Abstract. For an analysis of a molecular system from a computational statistical thermodynamics point of view, extensive molecular dynamics simulations are very inefficient. During this procedure, a lot of redundant data is generated. Whereas the algorithms spend most of the computing time for a sampling of configurations within the basins of the potential energy landscape of the molecular system, the important information about the long-time behaviour of the molecules is given by transition regions and barriers between the basins, which are sampled rarely only. Thinking of molecular dynamics trajectories, researchers try to figure out which kind of dynamical model is suitable for an efficient simulation. This article suggests to change the point of view from extensive simulation of molecular dynamics trajectories to more efficient sampling strategies of the conformation dynamics approach.

Keywords: metastability, conformation dynamics, rate matrix, molecular dynamics

PACS: 02.70.Lq, 02.70.Hm

INTRODUCTION

Classical molecular dynamics simulation (MD) is a widely used method for the analysis of molecular interactions. The equations of motion are solved on the basis of a given force-field which models the different mechanical aspects of covalent bonds and non-covalent interactions. Researchers try to estimate the essential behaviour of a molecular system by a statistical evaluation of the generated states. A state $x = (q, p)$ of a molecular system consisting of n atoms is given by a $6n$ -dimensional vector of $3n$ position coordinates $q \in \Omega$ and $3n$ momentum coordinates $p \in \Gamma$. The total energy $H(q, p)$ of a state is the sum of the kinetic energy $K(p)$ (only depending on the momentum coordinates) and the potential energy $V(q)$ (only depending on the position coordinates). Thus, H is separable. In MD, the equations of motion are solved:

$$\begin{aligned}\dot{q} &= \frac{\partial H}{\partial p} \\ \dot{p} &= -\frac{\partial H}{\partial q}.\end{aligned}\tag{1}$$

Standard MD simulation, however, would lead to incorrect statistical results if one wants to analyze a system at a given constant temperature T , with constant volume v and constant number of particles n . For this mvT -ensemble, Boltzmann derived the theoretically expected distribution of molecular states. The probability of a state x depends on the total energy $H(q, p)$ of this state. According to Boltzmann, the probability of a state x is proportional to

$$\pi(q, p) \propto \exp\left(-\frac{1}{k_B T} H(q, p)\right),\tag{2}$$

where T is the temperature and k_B the Boltzmann constant. In the followings we will always assume that the probability density functions can be normalized on a given position space Ω . In our case, H is separable, i.e. the Boltzmann distribution can be seen as the product distribution $\pi_q \cdot \pi_p$ of a distribution π_q of position coordinates and a distribution π_p of momentum variables. π_q and π_p are strictly positive functions. In this article, the separation of the density function into a position and a momentum part will be understood as follows: Independently from the position state q , the momentum variables are distributed according to π_p for each state $x = (q, p)$ of the system. Equation (2) shows an equilibrium distribution of a molecular ensemble, but it does not provide an equation of motion. Given an initial state x , how will the system evolve in time? Since Newton has formulated the equations of motion, we believe, that given a state x , we can predict the course of the molecular system. This course should somehow resemble (1). MD trajectories,

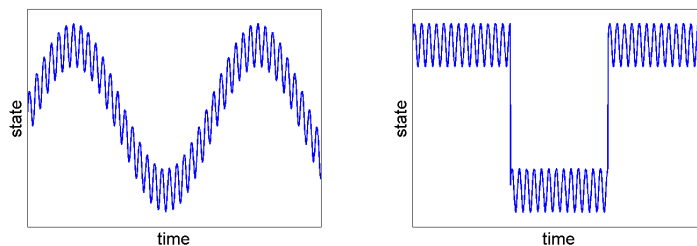


FIGURE 1. *Left.* A dynamical system including different timescales. *Right.* A flip-flop behaviour between two different metastable subsets in state space. The transitions between these sets are rare events. In contrast to the left curve, the transitions can not be described by a continuous dynamics on a different timescale sufficiently.

however, trace isolines of H . Hence, dynamical models have been invented which force the MD trajectories to change between the total energy levels according to the Boltzmann distribution (2). Many researchers have created models for a canonical ensemble dynamics, such that the distribution of simulation data of a single long-time trajectory converges to (2). They have been inspired by the equations of motion (1). Mainly two approaches are used in practise.

1. A deterministic approach: Instead of (1) an alternative similar deterministic dynamical system is defined which converges against Boltzmann distribution. A well-known example is the time-reversible Nosé-Hoover dynamics [1, 2]. Another example is the Berendsen thermostat [3] which does not generate the canonical ensemble exactly. Other time-reversible deterministic thermostats can be found in [4]. It should be mentioned, that the term “deterministic approach” is only of academic interest. From a numerical point of view, the Ljapunov exponent of the dynamical systems is usually very high: Long-time deterministic dynamical systems are chaotic. This is the reason why many researchers prefer molecular dynamics simulations for generating Boltzmann distributed ensembles.
2. A stochastic approach: Beside Smoluchowski [5] and Langevin dynamics [6], the class of hybrid Monte-Carlo methods (HMC) [7] is an example for a stochastic approach towards canonical ensemble dynamics. In HMC, the system is mainly propagated according to (1). Sole exception: After a certain time-span the momentum coordinates are refreshed randomly and a Metropolis-like acceptance step assures the convergence of the system towards (2). Since a total refreshing of momentum variables seems to be unphysical, there are alternative variants of this method. In these variants, momentum variables are more or less conserved, e.g., targeted shadow HMC [8].

The two approaches towards a canonical ensemble dynamics have an important property in common – the Markov property. Given a starting point $x = (q, p) \in \Omega \times \Gamma$, one can determine the probabilities for the possible future evolutions of the system. These probabilities only depend on the starting point x . From this point of view, a time-discretized computation of one of the mentioned dynamical models is nothing else but a realization of a Markov chain in phase space. In the above models, canonical ensemble dynamics try to combine the equation of motion (1) with a correct sampling of states according to (2). Beside possible physical inconsistencies of these models, there is always an unknown additional parameter which defines how fast the trajectories can change between the energy levels of H . From a physical point of view, this parameter determines the quality of the energy transfer of the molecular system with its environment in order to equilibrate temperature. This parameter is difficult to define and often appears arbitrarily. The aim of this article is to avoid analyzing single trajectories in phase space. From our point of view, the choice of a certain thermostated dynamics is not relevant. Instead of applying molecular dynamics, we in the following concentrate on conformation dynamics. The motivation is given by an observation: It is often claimed that molecular dynamical systems include many different timescales, starting from thermal vibration of bonds up to conformational changes of proteins or even larger systems. In our opinion, this statement might be misleading. A conformational change of a molecular system is a rare event, but the transition itself can be very fast, see Figure 1. Thus, on a larger timescale, the conformational changes of the molecular system can better be understood as a flip-flop behaviour [9]. From this flip-flop point of view, the dynamics is described by a matrix of transition probabilities (between “flip” and “flop”). The

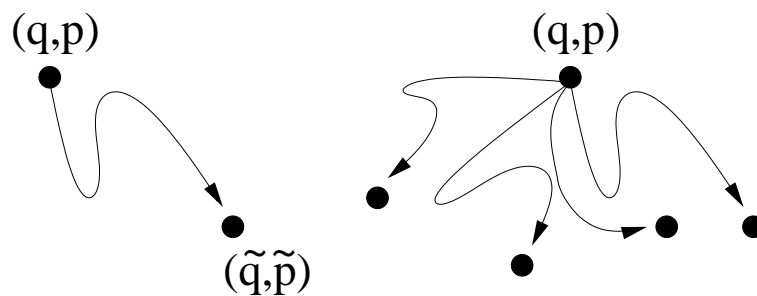


FIGURE 2. *Left.* In the case of (3), Hamiltonian dynamics is deterministic. A given initial state (q, p) leads to a fixed propagated state $(\tilde{q}, \tilde{p}) = \Phi_{-\tau}(q, p)$. *Right.* In the general case of (4) with a stochastic differential equation or with a Markov chain, the initial state (q, p) is propagated to different states with a different probability. $\Psi_{-\tau}(\cdot | (q, p))$ is the corresponding probability density function in Ω .

conformations of a molecular system are dynamically metastable subsets¹ of the *position space* [10]. For an analysis of rare events in canonical ensemble dynamics, one has to decompose the position space Ω into dynamically metastable subsets (conformations) and provide a transition probability matrix for the flip-flop behaviour between them². The paper aims at efficiency. It affirms that this analysis can be done without performing any dynamics simulations.

MARKOV PROPERTY AND SCHÜTTE OPERATOR \mathcal{T}

In (2), it is not claimed that a single ergodic long-time trajectory samples the phase space according to the Boltzmann distribution³. It is only claimed that a given Boltzmann distribution of initial states is conserved by the dynamical system. In this case, (1) may be a valid dynamical model preserving the canonical ensemble distribution. Note that Hamiltonian dynamics (1) is not ergodic in the canonical ensemble.

Conformations as metastable subsets of a molecular system are defined in position space only. Thus, the marginal distribution in Ω with regard to π_q will play an important role. Given a function $f : \Omega \rightarrow R$ in position space, the momentum-averaged effect of a dynamics simulation (1) on f is described by Schütte's operator $\mathcal{T}(\tau)$, see [10]:

$$\mathcal{T}(\tau)f(q) = \int_{\Gamma} f(\Pi_q \Phi^{-\tau}(q, p)) \pi_p(p) dp. \quad (3)$$

Equation (3) can be understood as follows: Given an initial state (q, p) , a backward Hamiltonian dynamics for a time-interval τ is performed. The new state is denoted as $\Phi^{-\tau}(q, p)$. Via Π_q , this new state is projected to position space. The integral in (3) averages over all possible initial momentum variables with given Boltzmann distribution π_p . In order to write down the operator (3), the Markov property of Hamiltonian dynamics is important. In the Hamiltonian case, the initial state (q, p) exactly determines the final state $\Phi^{-\tau}(q, p)$. Later on we will see that this definition of $\mathcal{T}(\tau)$ leads to an operator which is not time-harmonic. A more general definition of Schütte's operator using the Markov property is given by the *momentum-based transfer operator*:

$$\mathcal{P}(\tau)f(q) = \int_{\Gamma} \left(\int_{\Omega} f(\tilde{q}) \Psi_{-\tau}(\tilde{q} | (q, p)) d\tilde{q} \right) \pi_p(p) dp. \quad (4)$$

In equation (4), the initial state (q, p) determines a probability density function $\Psi_{-\tau}(\cdot | (q, p))$ for the possible evolutions of the system in position space. For an explanation see Figure 2. $\Psi_{-\tau}$ is a Dirac delta function in the case of a deterministic dynamics. Equation (4) can be used to define a momentum-based transfer operator for any of the dynamical models (deterministic and stochastic) mentioned above, even in the case of a dynamical model

¹ In this article, the meaning of the terms *conformation of a molecular system* and *metastable subset* is different from the common usage in chemistry. Both of the terms denote a subset of the configuration or position space Ω . The term *metastable* means that the subset is *almost stable* with regard to a dynamical process. In chemistry, a metastable subset can be understood as a basin of the potential energy function of the molecular system.

² We will see in the following that this is not possible in a strict sense.

³ Ergodicity is not needed, but from a physical point of view, only an ergodic dynamical model explains self-equilibration of molecular systems.

which is independent from momentum variables – like Smoluchowski dynamics [5]. Going from a trajectory based simulation of molecular systems according to one of the mentioned dynamical models to an analysis of the momentum-based transfer operator is the main step for gaining efficiency. This step leads us from analyzing single trajectories to analyzing a set of Boltzmann distributed trajectories. Instead of choosing a certain dynamical model, we will require certain physical properties for the operator $\mathcal{P}(\tau)$. In order to define \mathcal{P} , we already assumed the Markov property for the dynamical model. Furthermore, we will assume a detailed balanced Boltzmann equilibrium distribution and time-harmony for the operator \mathcal{P} in the following. In fact, these properties hold for more than one dynamical model Ψ_τ , see Appendix A.

STATE-BASED VERSUS ENSEMBLE-BASED TRANSITION PROBABILITIES

Hamiltonian dynamics is Markovian in phase space. It is not Markovian in position space. If we apply a projection and go from single states to sets of states in phase space, the Markov property is always lost. The same problem arises when we use the operator \mathcal{P} and a discretization of Ω in order to compute transition probabilities between subsets of Ω [11]. There is always a difference between *state-based transition probabilities* and the *fraction of states which go from set A to set B* in ensemble dynamics. In order to explain this difference, we assume a Markov chain in a finite state space. The transition probabilities of this chain can be expressed by a matrix P . In this example, the vector d is the invariant distribution of P with $d^\top P = d^\top$. Let χ_1, \dots, χ_k denote a set of characteristic vectors defining a decomposition of the state space into subsets. The element (i, j) of the dimension reduction P_c of P onto the given subsets can be written as:

$$P_c(i, j) = d_i^{-1} \chi_i^\top D P \chi_j,$$

where $D = \text{diag}(d)$ is the diagonal matrix of the vector d . The matrix P_c is stochastic: The row sums are 1 and the elements are non-negative. Note that the transition probability from a state q of the set χ_i to the set χ_j depends on q . It is not the same quantity for all $q \in \chi_i$. $P_c(i, j)$ can not be the correct probability, it is an average value. P_c is not the transition matrix of a Markov chain. See also [11] for a detailed example. Whereas, the transition probabilities for single states are different from the entries in P_c , $P_c(i, j)$ provides the correct *fraction of trajectories* going from i to j in one step of the Markov chain, if the initial states are distributed according to d projected to χ_i . This is an ensemble-based point of view. Let us now go back to continuous spaces. Due to the inherent Π_q -projection, \mathcal{P} is an ensemble-based transfer operator. Whenever we speak of *transition probabilities* derived from \mathcal{P} in the following, we think of the *fraction of states which go from A to B*. Conformation dynamics and molecular dynamics are substantially different research areas. In Conformation dynamics we do not aim at a realization of long-time trajectories. Note, however, that we always can formulate a Markov chain in position space on the basis of a given operator $\mathcal{P}(\tau)$. The procedure⁴ is as follows: Given a position state $q \in \Omega$, we randomly choose an initial momentum state $p \in \Gamma$ according to π_p . For this state $x = (q, p)$, we compute a realization of the dynamical model $\Psi_{-\tau}(\tilde{q}|(q, p))$ for the given time interval τ and end up with a new position state \tilde{q} . This algorithm is an *interpretation* of \mathcal{P} in terms of a Markov chain $q \rightarrow \tilde{q}$. For all mentioned dynamical models, this interpretation in fact provides an ergodic Markov chain with invariant density (2). The physical drawback is that this Markov chain is not interpretable as a trajectory, because it is discontinuous in momentum space after every time step τ . Therefore, it does not provide the “real” dynamics of the system expressed by \mathcal{P} and by its corresponding dynamical model $\Psi_{-\tau}$.

TIME-REVERSIBILITY AND SPECTRAL PROPERTIES OF \mathcal{P}

Hamiltonian dynamics (1) is time-reversible. Given an initial state $(q, p) \in \Omega \times \Gamma$ and its propagated state $(\tilde{q}, \tilde{p}) = \Phi^\tau(q, p)$, the following holds: $(q, -p) = \Phi^\tau(\tilde{q}, -\tilde{p})$. This fact is used in [10] to show $L_{\pi_q}^2(\Omega)$ -self-adjointness of Schütte’s operator $\mathcal{S}(\tau)$. An initial state (q, p) and its propagated state (\tilde{q}, \tilde{p}) have the same total energy H and, therefore, the same Boltzmann probability in equation (2), i.e. $\pi_q(\tilde{q}) \cdot \pi_p(\tilde{p}) = \pi_q(q) \cdot \pi_p(-p)$. In this equation $\pi_p(\tilde{p})$ denotes the probability for a transition $\tilde{q} \rightarrow q$ and $\pi_p(-p)$ denotes the probability for $q \rightarrow \tilde{q}$ (both backward in time).

⁴ We often show animations of molecular systems moving according to this kind of Markov chain in our presentations, but we always give a remark that this is a sampling of the position space, and we are only interested in the statistical results of the procedure, not in the special realization of the Markov chain.

Time reversibility can be written as $\pi_q(\tilde{q})P(\tilde{q} \rightarrow q) = \pi_q(q)P(q \rightarrow \tilde{q})$, where $P(a \rightarrow b)$ is the conditional probability for a transition from a to b . In order to determine transition probabilities for the transfer operator $\mathcal{P}(\tau)$, we have to take the mean transition probability with regard to the distribution of initial momentum variables p . In this general case, time-reversibility can be written as

$$\pi_q(\tilde{q}) \cdot \underbrace{\int_{\Gamma} \Psi_{-\tau}(q|\tilde{q}, p) \pi_p(p) dp}_{P(\tilde{q} \rightarrow q)} = \pi_q(q) \cdot \underbrace{\int_{\Gamma} \Psi_{-\tau}(\tilde{q}|(q, p)) \pi_p(p) dp}_{P(q \rightarrow \tilde{q})}. \quad (5)$$

Later on, we will replace Hamiltonian dynamics by another dynamical model. For this model, we will assume reversibility in the sense of (5). Equation (5) can be interpreted as *detailed balance condition* for the Boltzmann equilibrium distribution. This condition is required in the followings for the operator $\mathcal{P}(\tau)$. Given the $L^2_{\pi_q}(\Omega)$ scalar product $\langle \cdot, \cdot \rangle_{\pi_q}$, self-adjointness $\langle g, \mathcal{P}(\tau) f \rangle_{\pi_q} = \langle \mathcal{P}(\tau) g, f \rangle_{\pi_q}$ of $\mathcal{P}(\tau)$ in (4) is a direct consequence of equation (5):

$$\begin{aligned} \langle g, \mathcal{P}(\tau) f \rangle_{\pi_q} &= \int_{\Omega} g(q) \mathcal{P}(\tau) f(q) \pi_q(q) dq \\ &= \int_{\Omega} g(q) \int_{\Gamma} \int_{\Omega} f(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q, p)) d\tilde{q} \pi_p(p) dp \pi_q(q) dq \\ &= \int_{\Gamma} \int_{\Omega} \int_{\Omega} g(q) f(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q, p)) \pi_q(q) \pi_p(p) d\tilde{q} dq dp \\ &= \int_{\Gamma} \int_{\Omega} \int_{\Omega} g(q) f(\tilde{q}) \Psi_{-\tau}(q|\tilde{q}, p) \pi_q(\tilde{q}) \pi_p(p) dq d\tilde{q} dp \\ &= \int_{\Omega} \mathcal{P}(\tau) g(\tilde{q}) f(\tilde{q}) \pi_q(\tilde{q}) d\tilde{q} \\ &= \langle \mathcal{P}(\tau) g, f \rangle_{\pi_q}. \end{aligned} \quad (6)$$

Lemma 1. Given the operator $\mathcal{P}(\tau) : L^2_{\pi_q}(\Omega) \rightarrow L^2_{\pi_q}(\Omega)$ defined in (4) and the detailed balance condition (5), then $\mathcal{P}(\tau)$ is a linear, self-adjoint operator with $\|\mathcal{P}(\tau) f\|_{\pi_q} \leq \|f\|_{\pi_q}$. In particular, \mathcal{P} has a real-valued spectrum $\sigma(\mathcal{P})$ with $\sigma(\mathcal{P}) \subset [-1, 1]$.

Proof. Linearity is easy to check. Self-adjointness has been shown in equation (6). In order to prove $\|\mathcal{P}(\tau) f\|_{\pi_q} \leq \|f\|_{\pi_q}$, note that according to the Cauchy-Schwarz inequality, $(\int_{\Omega} f(x) \pi(x) dx)^2 \leq \int_{\Omega} f^2(x) \pi(x) dx$ for a probability density function $\pi : \Omega \rightarrow \mathbb{R}$. Thus, the following holds:

$$\begin{aligned} \|\mathcal{P}(\tau) f\|_{\pi_q}^2 &= \int_{\Omega} (\mathcal{P}(\tau) f(q))^2 \pi_q(q) dq \\ &= \int_{\Omega} \left(\int_{\Gamma} \int_{\Omega} f(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q, p)) d\tilde{q} \pi_p(p) dp \right)^2 \pi_q(q) dq \\ &\leq \int_{\Omega} \int_{\Gamma} \left(\int_{\Omega} f(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q, p)) d\tilde{q} \right)^2 \pi_q(q) \pi_p(p) dp dq \\ &\leq \int_{\Omega} \int_{\Gamma} \int_{\Omega} f^2(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q, p)) \pi_q(q) \pi_p(p) d\tilde{q} dp dq \\ (*) &= \int_{\Omega} \int_{\Gamma} \int_{\Omega} f^2(\tilde{q}) \Psi_{-\tau}(q|\tilde{q}, p) \pi_q(\tilde{q}) \pi_p(p) d\tilde{q} dp dq \\ &= \int_{\Omega} f^2(\tilde{q}) \pi_q(\tilde{q}) d\tilde{q} \\ &= \|f\|_{\pi_q}^2. \end{aligned}$$

The detailed balance condition has been used in (*). The spectral properties of \mathcal{P} are a consequence of self-adjointness and the above estimation. \square

IDENTIFICATION OF CONFORMATIONS

The inequality $\|\mathcal{P}(\tau) f\|_{\pi_q} \leq \|f\|_{\pi_q}$ is sharp. It will be shown that certain indicator functions f_A are eigenfunctions of \mathcal{P} corresponding to the eigenvalue $\lambda = 1$. *Stable subsets* $A \subset \Omega$ of Ψ_{τ} are defined as follows: If a position state q is in the closure of A , then $\int_{\Gamma} \Psi_{-\tau}(\tilde{q}|(q, p)) \pi_p(p) dp = 0$ for all $\tilde{q} \notin A$ and $\tau \in \mathbb{R}$. If $A \subset \Omega$ is a stable subset, then its complement $\Omega \setminus A$ is also stable. This can be shown by applying the detailed balance condition (5) for $\tilde{q} \notin A$ and $q \in A$. The definition of stability means that there is no transition between A and its complement $\Omega \setminus A$.

Lemma 2. If $f_A : \Omega \rightarrow \{0, 1\}$ is the indicator function of a stable subset $A \subset \Omega$, then $\mathcal{P}(\tau) f_A = f_A$.

Proof. Stability of A and $\Omega \setminus A$ directly implies $\mathcal{P}(\tau) f_A(q) = 0$ for all $q \notin A$, because the integrand $f_A(\tilde{q}) \int_{\Gamma} \Psi_{-\tau}(\tilde{q}|(q, p)) \pi_p(p) dp$ is zero for all $\tilde{q} \in \Omega$. For $q \in A$:

$$\begin{aligned} \mathcal{P}(\tau) f_A(q) &= \int_{\Gamma} \left(\int_{\Omega} f_A(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q, p)) d\tilde{q} \right) \pi_p(p) dp \\ &= \int_{\Gamma} \left(\int_A f_A(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q, p)) d\tilde{q} \right) \pi_p(p) dp \\ &= \int_{\Gamma} \left(\int_A \Psi_{-\tau}(\tilde{q}|(q, p)) d\tilde{q} \right) \pi_p(p) dp \\ &= 1. \end{aligned}$$

The last equality uses the fact, that $\Psi_{-\tau}$ is a density function that is normalized in A . \square

There are different possibilities to define *metastable subsets* (conformations) of the dynamical model on the basis of the above Lemma. One possibility is to define metastable subsets by indicator functions f , which meet $\mathcal{P}(\tau) f \approx f$. This approach has been investigated for many years for the Schütte operator (3) by the groups of Deuffhard and Schütte [10, 12, 13, 5]. A novel approach [14, 15, 16] towards conformation dynamics replaces indicator functions $f : \Omega \rightarrow \{0, 1\}$ by almost characteristic functions (fuzzy sets) $\chi : \Omega \rightarrow [0, 1]$. In this approach, a conformation-based decomposition of Ω is given by a set of k metastable almost characteristic functions $\chi_1, \dots, \chi_k : \Omega \rightarrow [0, 1]$. The functions meet the partition-of-unity property, i.e. for all $q \in \Omega$ the membership values have the sum $\sum_{i=1}^k \chi_i(q) = 1$. In practice, the first step of the identification of conformations is an $n \times n$ -Galerkin discretization $P(\tau)$ of $\mathcal{P}(\tau)$:

$$P(\tau) = (\langle \varphi, \varphi \rangle_{\pi_q})^{-1} \langle \varphi, \mathcal{P}(\tau) \varphi \rangle_{\pi_q},$$

where $\varphi = \{\varphi_1, \dots, \varphi_n\}$ is a set of $n \gg k$ indicator functions $\varphi_i : \Omega \rightarrow \{0, 1\}$ of a decomposition of Ω . The discretized matrix $P(\tau)$ is generalized symmetric because $\mathcal{P}(\tau)$ is π_q -self-adjoint. More precisely, $DP(\tau)$ is symmetric, where $D = \text{diag}(d_1, \dots, d_n) = \langle \varphi, \varphi \rangle_{\pi_q}$ is the diagonal matrix of weights⁵ $d_i := \int_{\Omega} \varphi_i(q) \pi_q(q) dq$. Therefore, $P(\tau)$ has a real-valued spectrum $\sigma(P(\tau)) \subset [-1, 1]$ with D -orthonormal eigenvectors. $P(\tau)$ is a stochastic matrix. Robust Perron Cluster Analysis (PCCA+, [15]) can be used to compute the membership functions χ on the basis of this matrix. The membership functions are given as linear combination $\chi = \varphi C$ of the indicator functions, with an $n \times k$ -transformation matrix C . This matrix $C = X \mathcal{A}$ is determined via PCCA+ as a linear combination of the k dominant eigenvectors $X = \{X_1, \dots, X_k\}$ of $P(\tau)$. PCCA+ aims at an optimal linear transformation $\mathcal{A} \in \mathbb{R}^{k \times k}$ such that χ meets the above conditions (non-negativity and partition-of-unity). With $\mathcal{X} := \varphi X$, the computation of membership functions can be written as $\chi = \mathcal{X} \mathcal{A}$, where \mathcal{A} is the result of the PCCA+ algorithm.

Assumption. At this stage, we give the main assumption of our approach towards an effective analysis of rare events. This assumption can not be shown mathematically. We will assume, that via spectral analysis of $P(\tau)$, we can compute $\mathcal{X} = \varphi X$ as a good approximation of a set of eigenfunctions of $\mathcal{P}(\tau)$ and their corresponding eigenvalues. This assumption has two implications:

1. $\mathcal{P}(\tau)$ has a set of eigenfunctions. For the Schütte operator $\mathcal{S}(\tau)$, conditions have been formulated for which \mathcal{S} has a set of discrete dominant eigenvalues. These estimations are based on a perturbation analysis of Lemma 2.
2. The decomposition φ of Ω allows for a good approximation of the eigenfunctions of $\mathcal{P}(\tau)$. The dominant eigenfunctions of $\mathcal{P}(\tau)$ (if they exist) are almost constant within the metastable sets. Major “gradients” of the eigenfunctions can be found in transition regions. Therefore, an adequate decomposition of Ω is fine enough in these transition regions.

⁵ Note that for indicator functions $\varphi_i^2 = \varphi_i$ and $\varphi_i \varphi_j = 0$ for $i \neq j$.

ANALYSIS OF RARE EVENTS

The element $P_{ij}(\tau)$ of the i -th column and the j -th row of $P(\tau)$ is

$$P_{ij}(\tau) = \frac{\int_{\Omega} \varphi_i(q) \int_{\Gamma} \left(\int_{\Omega} \varphi_j(\tilde{q}) \Psi_{-\tau}(\tilde{q}|(q,p)) d\tilde{q} \right) \pi_p(p) \pi_q(q) dp dq}{d_i},$$

where the numerator provides the transitions between φ_i and φ_j for Boltzmann distributed initial states. The denominator gives the statistical weight d_i of the discretization set φ_i . $P_{ij}(\tau)$ provides the fraction of trajectories going from φ_i to φ_j in time τ . Due to reversibility (5), the discrete detailed balance condition $d_i \cdot P_{ij}(\tau) = d_j \cdot P_{ji}(\tau)$ holds. From an ensemble-based point of view, we will denote P as transition matrix. In the case of almost characteristic functions χ , the matrix

$$P_c(\tau) := (\langle \chi, \chi \rangle_{\pi_q})^{-1} \langle \chi, \mathcal{P}(\tau) \chi \rangle_{\pi_q},$$

can be seen as transition matrix between the conformations χ , analogously. The functions χ and this matrix P_c are the desired results of conformation dynamics. In order to extract the essential long time behaviour of the molecular system, conformation dynamics aims at an identification of conformations with maximal life-times (low transition probabilities between them). An optimal choice of χ maximizes the trace of $P_c(\tau)$, i.e. the probability to stay in a conformation. We assume that there is a set \mathcal{X} of k eigenfunctions of \mathcal{P} with an $k \times k$ -diagonal matrix Λ of eigenvalues. In the case of PCCA+ with $\chi = \mathcal{X} \mathcal{A}$, we have

$$\begin{aligned} P_c(\tau) &= (\langle \chi, \chi \rangle_{\pi_q})^{-1} \langle \chi, \mathcal{P}(\tau) \chi \rangle_{\pi_q} \\ &= (\mathcal{A}^\top \mathcal{A})^{-1} (\mathcal{A}^\top \Lambda \mathcal{A}) \\ &= \mathcal{A}^{-1} \Lambda \mathcal{A}, \end{aligned}$$

where we used the fact, that \mathcal{X} is a π_q -orthogonal set of eigenfunctions due to Lemma 1 and \mathcal{X} can be normalized according to the π_q -scalar product. The trace of P_c equals the trace of Λ , i.e. the sum of the eigenvalues of the selected eigenfunctions. The choice of χ is optimal if the selected eigenfunctions correspond to the dominant eigenvalues of \mathcal{P} .

TIME-HARMONY AND INFINITESIMAL GENERATOR

Time-harmony is the next physical aspect we require for the operator \mathcal{P} . From a physical point of view, this property means that the heat bath is continuously and steadily acting on the molecular system. From a mathematical point of view, $\mathcal{P}(\tau)$ should meet special continuity conditions: We require that, for all $f \in L^2_{\pi_q}(\Omega)$, the following limit exists

$$\mathcal{Q}f := \lim_{\tau \rightarrow 0} \frac{\mathcal{P}(\tau)f - f}{\tau}. \quad (7)$$

Equation (7) defines an operator \mathcal{Q} . Furthermore, the operator \mathcal{P} should be time-harmonic, which means $\mathcal{P}(t+s) = \mathcal{P}(t)\mathcal{P}(s)$. In this situation, the operator \mathcal{Q} is the infinitesimal generator of the half-group \mathcal{P} of operators:

$$\mathcal{P}(\tau) = \exp(\tau \mathcal{Q}).$$

The latter condition is not valid for the Schütte operator $\mathcal{T}(\tau)$. In this case⁶, $\mathcal{T}(t+s) \neq \mathcal{T}(t)\mathcal{T}(s)$, see [5]. Therefore, we define the more general operator (4) and require time-harmony. The existence of at least two dynamical models which meet self-adjointness and time-harmony is shown in Appendix A. The existence of an infinitesimal generator \mathcal{Q} of \mathcal{P} leads to the definition of the Galerkin discretization Q of \mathcal{Q} :

$$Q := (\langle \varphi, \varphi \rangle_{\pi_q})^{-1} \langle \varphi, \mathcal{Q} \varphi \rangle_{\pi_q}.$$

⁶ The Markov chain which can be associated with Schütte's operator $\mathcal{T}(\tau)$ is a Hamiltonian dynamics with a heat bath contact, i.e. a refresh of momenta, at discrete points in time $0, \tau, 2\tau, 3\tau, \dots$

Using equation (7), we can conclude:

$$Q = \lim_{\tau \rightarrow 0} \frac{P(\tau) - I}{\tau}, \quad (8)$$

where I is the identity matrix. In (8), the elements of the Q -matrix measure the transition rates between the discretization sets of Ω (in an ensemble-based sense). Equation (8) is used in the next section to compute Q . Although equation (8) and $\mathcal{P}(t+s) = \mathcal{P}(t)\mathcal{P}(s)$ hold, Q is not an infinitesimal generator of $P(\tau)$ in general. The time-harmony might not be valid on the level of the Galerkin discretization: $P(t+s) \neq P(t)P(s)$. Fortunately, the discretization step is just an intermediate between the continuous operator $\mathcal{P}(\tau)$ and the transition matrix $P_c(\tau)$. The following Lemma is important:

Lemma 3. If $\mathcal{X} = \{\mathcal{X}_1, \dots, \mathcal{X}_k\}$ is a π_q -orthonormal set of k eigenfunctions of $\mathcal{P}(\tau)$ and $\chi = \mathcal{X}\mathcal{A}$ a regular basis transformation of these eigenfunctions then

$$\mathcal{P}(\tau) = \exp(\tau\mathcal{Q}) \quad \Rightarrow \quad P_c(\tau) = \exp(\tau Q_c),$$

where $P_c(\tau) = (\langle \chi, \chi \rangle_{\pi_q})^{-1} \langle \chi, \mathcal{P}(\tau)\chi \rangle_{\pi_q}$ and $Q_c = (\langle \chi, \chi \rangle_{\pi_q})^{-1} \langle \chi, \mathcal{Q}\chi \rangle_{\pi_q}$.

Proof. The proof is a simple calculation. If Λ is the diagonal matrix of the eigenvalues of $\mathcal{P}(\tau)$ with regard to the eigenfunctions \mathcal{X} and M is the diagonal matrix of the eigenvalues of \mathcal{Q} (with the same eigenfunctions \mathcal{X}), then $\Lambda = \exp(\tau M)$. Furthermore, $P_c(\tau) = \mathcal{A}^{-1}\Lambda\mathcal{A}$ and $Q_c = \mathcal{A}^{-1}M\mathcal{A}$. \square

The above Lemma shows that (assuming the existence of eigenfunctions of \mathcal{P}) Robust Perron Cluster Analysis $\chi = \mathcal{X}\mathcal{A}$ is the correct way of dimension reduction. If the set φ of indicator functions allows for a good approximation of the dominant eigenfunctions $\mathcal{X} = \varphi X$ of \mathcal{P} , where X is the set of dominant eigenfunctions of $P(\tau)$, then $Q_c = \mathcal{A}^{-1}M\mathcal{A}$ is the infinitesimal generator of P_c according to the above Lemma. Therefore, in the next section, we will assume that φ is an adequate set of indicator functions. We will approximate Q and compute the transformation matrix \mathcal{A} via PCCA+ on the basis of the dominant eigenvectors X of Q . The desired transition matrix⁷ is

$$P_c(\tau) = \exp(\tau Q_c) = \mathcal{A}^{-1} \exp(\tau M) \mathcal{A}, \quad (9)$$

where $M \in R^{k \times k}$ is the diagonal matrix of the dominant eigenvalues of Q .

APPROXIMATION OF Q

Using equation (9), an approximation of a discretized infinitesimal generator Q provides all information for the computation of P_c : the transformation matrix \mathcal{A} (via PCCA+ based on the eigenvectors X of Q) and the matrix M of its dominant eigenvalues. Q is independent from time and from the current state of the system. It is not relevant, if we assume an equilibrated system or a system on its way to equilibration, because it is always the same dynamical process acting on the molecular system defined by Q . In the following, the approximation of Q is based on an analysis of an equilibrated system. Q corresponds to a matrix of reaction rates. According to (8), the outer diagonal elements of Q_{ij} measure the transition rates between the discretization boxes φ_i and φ_j with regard to the ‘‘concentration’’ d_i of states in φ_i . For the reason of mass conservation, the diagonal elements of Q equal the negative row-sum of its outer diagonal elements, see also [17]. Via normalization of the rows, Q can be written in the form:

$$Q = R(K - I), \quad (10)$$

where I is the identity matrix, K is a stochastic matrix, with $K_{ii} = 0$ for all diagonal elements $i = 1, \dots, n$, and R is a diagonal matrix of rate factors. An instantaneous transition is only possible between neighboring discretization boxes in Ω , because the physical motivation for the definition of (4) and for Figure 2 is a dynamical system, continuous in time and space. Only the states which cross the intersection S_{ij} between φ_i and φ_j contribute to the transition rate

⁷ P_c is not a transition matrix in a strict sense. P_c is not based on sets but on membership functions χ . This means, that conformations are not subsets of Ω , they are ‘‘fuzzy’’-subsets.

between φ_i and φ_j . The states located on the intersection S_{ij} are weighted according to the Boltzmann distribution. For each row i of K , the transition rates are proportional to a surface integral

$$K_{ij} \propto \int_{S_{ij}} w_{ij}(q) \pi_q(q) dq, \quad (11)$$

where $w_{ij}(q)$ denotes the intensity of the flux from φ_i to φ_j through the position state q . According to Gauss' theorem, the intensity $w_{ij}(q)$ of the flux between φ_i and φ_j is given by the scalar product of the time derivative \dot{q} and the outgoing normal vector $n_{ij}(q)$ of the surface S_{ij} in q . Instead of one time derivative \dot{q} at position q , we have an ensemble of trajectories going through q . We assume, that the dynamics locally behaves like (1). In this case,

$$\begin{aligned} w_{ij}(q) &= \int_{\Gamma} \max \left\{ \frac{\partial H(q, p)}{\partial p} \cdot n_{ij}(q), 0 \right\} \pi_p(p) dp \\ &= \int_{\Gamma} \max \{ \mathbf{M}^{-1} p \cdot n_{ij}(q), 0 \} \pi_p(p) dp, \end{aligned} \quad (12)$$

where \mathbf{M} is the diagonal matrix of atom masses of the observed molecule. The computation of Q is like in [11]. Briefly: Because of the normalization of K , only the computation of ratios K_{ij}/K_{il} is needed for the determination of K . We will see in the next section that it is possible to define a Markov chain between the surfaces S_{ij} and S_{il} in order to compute these ratios via Monte Carlo quadrature numerically. A simple calculation shows that $d^\top Q = 0$ for the weights d_1, \dots, d_n of the discretization boxes. This can be used for the computation of the rate factors $R = \text{diag}(r_1, \dots, r_n)$ in (10). $d^\top Q = 0$ is equivalent to the linear system $r^\top D(K - I) = 0$ with $D = \text{diag}(d_1, \dots, d_n)$. Therefore, the rate factors $r^\top = (r_1, \dots, r_n)$ can be computed except for an unknown common scaling factor μ which is called *timescale factor*. In order to compute the weights d_1, \dots, d_n of the discretization boxes, we will define a Markov chain in the next section for a Monte Carlo quadrature of

$$d_i \propto \int_{\Omega} \varphi_i(q) \pi_q(q) dq. \quad (13)$$

In the presented approach, the computation of statistical weights d is done *before* Q is estimated. This is a further advantage of the presented approach. Usually, in molecular dynamics, the statistical weights are estimated *on the basis of* transition rates or transition probabilities derived from the simulation data. Note, however, that a computation of statistical weights based on transition probabilities without aggregation/disaggregation strategies is ill-conditioned [18].

MARKOV CHAIN MONTE-CARLO-QUADRATURE

In this section, we just give some algorithmic ideas for the computation of the integrals (11) and (13). The software development of routines for realistic molecules can be based on these ideas. The result of conformation dynamics is an $k \times k$ transition matrix P_c together with k conformations given by almost characteristic functions $\chi_1, \dots, \chi_k : \Omega \rightarrow [0, 1]$. Instead of a statistical analysis of long-time trajectories for the computation of these quantities, this article provides a method for the estimation of the infinitesimal generator Q_c of P_c . The estimate can be done by evaluating the integrals (11) and (13) and by discretizing the position space into subsets $\varphi_1, \dots, \varphi_n$ in such a way, that eigenfunctions of \mathcal{P} can be approximated sufficiently by step functions based on the decomposition φ . The second condition for a good estimate of Q_c can be assured by an hierarchical adaptive procedure: For a given decomposition φ we compute the conformations χ . Afterwards, we identify two sets φ_i and φ_j , such that χ varies at most between these two sets. For the next step, we hierarchically refine φ in this region and go on from the beginning. For a two dimensional example, the result φ of an adaptive refinement procedure can look like Fig. 4, right side. The black lines in Fig. 4 correspond to the intersecting planes S_{ij} for (11), whereas the boxes represent the discretization sets φ for (13). The integrals (11) and (13) are of the general form

$$\int_{A_i} f_i(q) \pi_q(q) dq.$$

In (11), the sets A_i are given by the intersection planes S_{ij} and $f_i = w_{ij}$. In (13), the sets A_i are given by the discretization sets φ_i and $f_i \equiv 1$ is a constant function. Since the row sum of K is one and also the sum of the weights d_i is one, it is sufficient to compute ratios of these integrals

$$r_{ij} := \frac{\int_{A_i} f_i(q) \pi_q(q) dq}{\int_{A_j} f_j(q) \pi_q(q) dq} = \frac{\int_{A_i} f_i(q) \exp(-\beta V(q)) dq}{\int_{A_j} f_j(q) \exp(-\beta V(q)) dq}.$$

We can expand these ratios into three terms

$$r_{ij} := \underbrace{\frac{\exp(-\beta V(q_i^*))}{\exp(-\beta V(q_j^*))}}_I \cdot \underbrace{\frac{\int_{A_i} \exp(-\beta(V(q) - V(q_i^*))) dq}{\int_{A_j} \exp(-\beta(V(q) - V(q_j^*))) dq}}_{II} \cdot \underbrace{\frac{\int_{A_i} f_i(q) \frac{\pi_q(q)}{\int_{A_i} \pi_q(\tilde{q}) d\tilde{q}} dq}{\int_{A_j} f_j(q) \frac{\pi_q(q)}{\int_{A_j} \pi_q(\tilde{q}) d\tilde{q}} dq}}_{III}, \quad (14)$$

where q_i^* might be a minimum of V in A_i . Term I in (14) is used for shifting potential energies in order to make term II numerically tractable. Term I can be evaluated directly. Term III in (14) is equal to 1 for (13). It can be computed by Metropolis Monte Carlo algorithms in (11). Metropolis Monte Carlo is used to sample from the restricted Boltzmann distribution in A_i and the integrals are approximated by an average of f_i over all sampled position states. Term II in (14) can be estimated by applying a jump algorithm (similar to ConfJump in [19]). The jump algorithm can be organized in such a way that it provides an estimation of term II and a sampling for term III at the same time. The corresponding algorithm is a Metropolis Monte Carlo sampling within the sets A_i and A_j combined with a jump step between the sets A_i and A_j : During a Metropolis Monte Carlo sampling of one set, say A_i , with a certain probability a *jump* to another set, say A_j , is proposed. The starting point $\tilde{q} \in A_j$ is sampled according to an artificial but explicitly given density function $h_j : A_j \rightarrow R$. In accordance with the Metropolis Monte Carlo sampling scheme, this jump step is accepted with probability

$$p_{acc} = \min \left\{ \frac{h_i(q) \cdot \exp(-\beta(V(\tilde{q}) - V(q_i^*)))}{h_j(\tilde{q}) \cdot \exp(-\beta(V(q) - V(q_i^*)))}, 1 \right\},$$

where $q \in A_i$ is the last sampled point in A_i . If the jump step is accepted, the Metropolis Monte Carlo sampling is continued in A_j with starting point \tilde{q} . If the jump step is rejected, the Metropolis Monte Carlo sampling continues in set A_i with starting point q . At the end of the overall sampling procedure, the sampling points in A_i and A_j are counted. The ratio II is given by the ratio of these numbers. The shift in term I is introduced to equilibrate the jumps between A_i and A_j . The jump approach for the evaluation of (14) can be generalized from two domains A_i, A_j to multiple domains A_1, \dots, A_k . The only missing quantity is an evaluation of the integrand w_{ij} in term III at the generated sampling points. Given a sampling $q^{(1)}, \dots, q^{(M)} \in S_{ij}$ of π_q in an intersecting hyperplane S_{ij} , the weight $w_{ij}(q^{(m)})$ of the flux between i and j through $q^{(m)}$ can be estimated by taking the average of $\max\{\mathbf{M}^{-1} p \cdot n_{ij}(q^{(k)}), 0\}$ for a sampling of momentum vectors p according to π_p .

ILLUSTRATIVE EXAMPLE

In order to give an illustrative example how efficiency of the analysis of rare events is increased by the approach proposed in this article, the following two-dimensional potential energy function is investigated:

$$\begin{aligned} V(x, y) = & 3 \exp\left(-x^2 - \left(y - \frac{1}{3}\right)^2\right) - 3 \exp\left(-x^2 - \left(y - \frac{5}{3}\right)^2\right) \\ & - 5 \exp\left(-(x-1)^2 - y^2\right) - 5 \exp\left(-(x+1)^2 - y^2\right) \\ & + 0.2x^4 + 0.2\left(y - \frac{1}{3}\right)^4. \end{aligned} \quad (15)$$

A contour plot of this x -symmetric function is shown in Fig. 3, see also [20]. The potential energy function has three local minima. If we start a long-time trajectory-based sampling of this potential energy function, the procedure will be trapped in one of the local minima, see Fig. 4, left side. We start our adaptive hierarchical approach for an analysis of this system with a decomposition of the position space Ω into 8×8 discretization boxes φ . At the end of our procedure, the final discretization looks like in Fig. 4, right side. The algorithm spends most of the time with refining the grid in transition regions, where the important information about transition probabilities is given. For the computation of the membership functions χ , Q -matrices have been estimated using Metropolis Monte Carlo samplings combined with jump steps on the basis of equation (14). In the present example, term III in equation (14) is 1 for the evaluation of (11), because the masses \mathbf{M} are equal 1 (radial symmetry of the π_p -distribution) and the intersection planes are along the q -coordinates of the system. Only term I and term II are left for computation. The final result is a generator Q_c for

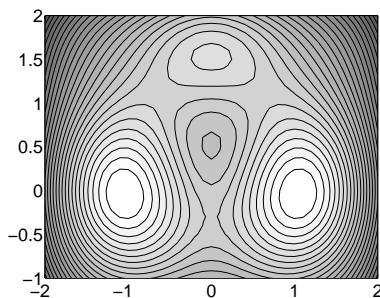


FIGURE 3. Potential energy function. The function has three local minima. Two deeper ones at about $(-1;0)$, $(1;0)$, and one minimum at about $(0;1.5)$.

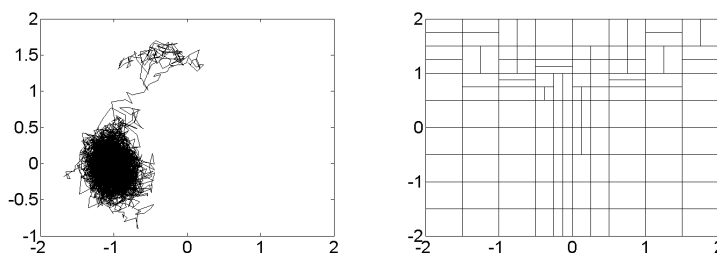


FIGURE 4. *Left.* Realization of a reversible Markov chain starting in the upper minimum region of (15). The Markov chain is trapped in one of the deeper minima. The simulation does not provide transition probabilities. *Right.* Final discretization of Ω for an adaptive approximation of almost characteristic membership functions χ . In contrast to the realization of a Markov chain, the discretization is refined in transition regions, where the important information about transition probabilities is “located”.

a decomposition of the position space into 3 conformations (the three local minima of V , see Fig. 5):

$$Q_c = \begin{pmatrix} -0.001744 & 0.000840 & 0.000904 \\ 0.000004 & -0.000005 & 0.000001 \\ 0.000004 & 0.000001 & -0.000005 \end{pmatrix}.$$

This matrix shows that the upper local minimum of V (first row of Q_c) is much less stable than the other two minima. The symmetry of the two deeper minima results in a certain accordance of the rates in Q_c . This example clearly shows the main problem of long-time trajectories in molecular simulations. Molecular systems consist of many metastable subsets of Ω having very different life times. A timescale free description of the system is mandatory. In contrast to $P_c(\tau)$, Q_c does not depend on a given time interval τ .

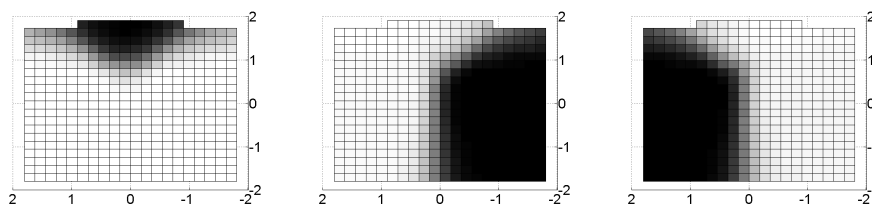


FIGURE 5. Membership functions χ of the three metastable parts of the potential in Fig. 3 (white=0, black=1). The membership functions are nearly constant inside the metastable parts. High gradients can be found in transition regions where the refinement is done, see Fig. 4, right side.

CONCLUSION

In this article, the conformation dynamics approach has been investigated. It has been shown that certain physical properties of a canonical ensemble dynamics lead to the definition of an infinitesimal generator \mathcal{Q} which describes the dynamics of the molecular system. This generator can be used to derive the essential dynamics on the basis of a low-dimensional transition matrix $P_c(\tau)$ between metastable parts of the position space. A numerical approximation of a discretized operator Q can be derived without computing long-time trajectories. Instead of the simulation of trapped long-time trajectories leading to bad statistics, a Monte-Carlo quadrature in high-dimensional spaces is used. Efficiency for the computation of Q is gained by focussing the Monte-Carlo quadrature points in transition regions of Ω where the important information about transition probabilities is located.

APPENDIX A: EXISTENCE OF A DYNAMICAL MODEL

Smoluchowski dynamics is an example for a possible dynamical model which meets time-reversibility of Ψ_τ and time-harmony of \mathcal{P} , see [5]. This type of dynamics is valid for our considerations. Smoluchowski dynamics is a *first* order stochastic differential equation in Ω . Therefore Ψ_τ is independent from the momentum variables p . The given requirements (condition (5) and time-harmony), however, are not solely valid for dynamical systems which only depend on position coordinates: Assume for some $\tau > 0$, there is a self-adjoint operator $\mathcal{P}(\tau)$ with a coupled system of (stochastic) differential equations. One example is Schütte's operator $\mathcal{T}(\tau)$ for Hamiltonian dynamics. The product of two momentum-based transfer operators $\mathcal{P}_3(\tau) := \mathcal{P}_1(\tau)\mathcal{P}_2(\tau)$ is again a transfer operator by showing that

$$\Psi_{-\tau}^{(3)}(\tilde{q}|(q, p)) = \int_{\Omega} \int_{\Gamma} \Psi_{-\tau}^{(1)}(\tilde{q}|(\tilde{q}, \tilde{p})) \cdot \Psi_{-\tau}^{(2)}(\tilde{q}|(q, p)) \pi_p(\tilde{p}) d\tilde{p} d\tilde{q},$$

is a density function. Using this fact, the expression $\exp(s\mathcal{T}(\tau))$ can be written in terms of (4). An adequate normalization

$$\overline{\mathcal{P}}(s) := \frac{\exp(s\mathcal{T}(\tau))}{\exp(s)}$$

defines a transfer operator $\overline{\mathcal{P}}$. Since $\overline{\mathcal{P}}(s) = \exp(s(\mathcal{T}(\tau) - \mathcal{I}))$ with the identity operator \mathcal{I} , $\overline{\mathcal{D}} := \mathcal{T}(\tau) - \mathcal{I}$ is the infinitesimal generator of $\overline{\mathcal{P}}$. The generator $\overline{\mathcal{D}}$ and the generated operators are self-adjoint. $\overline{\mathcal{P}}$ is time-harmonic. Every eigenfunction of $\mathcal{T}(\tau)$ with eigenvalue λ is an eigenfunction of $\overline{\mathcal{P}}(s)$ with eigenvalue $\exp(s(\lambda - 1))$. Especially, indicator functions f_A of stable subsets $A \subset \Omega$ are eigenfunctions of $\overline{\mathcal{P}}$ with $\overline{\mathcal{P}}(s)f_A = f_A$. A realization of the dynamical model $\Psi_s(\cdot|(q, p))$ of $\overline{\mathcal{P}}(s)$ can be done as follows: For $s > 0$ and an initial state (q, p) select $\Phi^k(\tau)(q, p)$ as next step of the dynamical model with a probability of $\frac{s^k}{k! \exp(s)}$. For $s = 0$, the identity ($k = 0$) is chosen. For increasing s , higher k values are more and more probable. Note that this procedure is not independent from p . This is only an artificial example, because the "time" variable s does not have the physical meaning of time.

REFERENCES

1. D. Frenkel, and B. Smit, *Understanding Molecular Simulation – From Algorithms to Applications*, vol. 1 of *Computational Science Series*, Academic Press, 2002.
2. W. G. Hoover, *Time Reversibility, Computer Simulation and Chaos*, World Scientific, 1999.
3. H. J. C. Berendsen, J. P. M. Postma, W. F. V. Gunsteren, A. Dinola, and J. R. Haak, *Journal of Chemical Physics* **81**, 3684–3690 (1984).
4. W. G. Hoover, A. Kenichiro, C. G. Hoover, and S. V. de Groot, *Physica D: Nonlinear Phenomena* **187**, 253–267 (2004).
5. W. Huisinga, *Metastability of Markovian systems: A transfer operator based approach in application to molecular dynamics*, Doctoral thesis, Freie Universität Berlin (2001).
6. T. Schlick, *Molecular Modelling and Simulation*, Springer, 2002.
7. S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Phys. Lett. B* **195**, 216–222 (1987).
8. E. Akhmatkaya, and S. Reich, "The Targeted Shadowing Hybrid Monte Carlo (TSHMC) Method," in *New Algorithms for Macromolecular Simulation, Lecture Notes in Computational Science and Engineering*, edited by B. Leimkuhler, C. Chipot, R. Elber, A. Laaksonen, A. Mark, T. Schlick, C. Schütte, and R. Skeel, 2006, vol. 49, pp. 145–158.
9. P. Deuffhard, and C. Schütte, "Molecular Conformation Dynamics and Computational Drug Design," in *Applied Mathematics Entering the 21st Century. Proceedings ICIAM 2003*, edited by J. M. Hill, and R. Moore, 2004, pp. 91–119, invited paper.

10. C. Schütte, *Conformational Dynamics: Modelling, Theory, Algorithm, and Application to Biomolecules*, Habilitation thesis, Department of Mathematics and Computer Science, Freie Universität Berlin (1999).
11. M. Weber, Conformation-based transition state theory, ZIB Report 07-18, Zuse Institute Berlin (2007).
12. P. Deuffhard, W. Huisinga, A. Fischer, and C. Schütte, *Lin. Alg. Appl.* **315**, 39–59 (2000).
13. A. Fischer, C. Schütte, P. Deuffhard, and F. Cordes, “Hierarchical Uncoupling–Coupling of Metastable Conformations,” in *Computational Methods for Macromolecules: Challenges and Applications – Proc. of the 3rd Intern. Workshop on Algorithms for Macromolecular Modelling*, edited by T. Schlick, and H. H. Gan, Springer, Berlin, Heidelberg, New York, 2002, pp. 235–259.
14. M. Weber, and T. Galliat, Characterization of Transition States in Conformational Dynamics using Fuzzy Sets, ZIB-Report 02-12, Zuse Institute Berlin (2002).
15. P. Deuffhard, and M. Weber, “Robust Perron Cluster Analysis in Conformation Dynamics,” in *Lin. Alg. App. – Special Issue on Matrices and Mathematical Biology*, edited by M. Dellnitz, S. Kirkland, M. Neumann, and C. Schütte, Elsevier, 2005, vol. 398C, pp. 161–184.
16. M. Weber, *Meshless Methods in Conformation Dynamics*, Doctoral thesis, Department of Mathematics and Computer Science, Freie Universität Berlin (2006), published by Verlag Dr. Hut, München.
17. M. Kijima, *Markov Processes for Stochastic Modeling*, Stochastic Modeling Series, Chapman and Hall, 1997.
18. M. Weber, S. Kube, L. Walter, and P. Deuffhard, *SIAM J. Multisc. Mod.* **6**, 396–416 (2007).
19. L. Walter, and M. Weber, ConfJump: a fast biomolecular sampling method which drills tunnels through high mountains, ZIB-Report 06-26, Zuse Institute Berlin (2006).
20. P. Metzner, C. Schütte, and E. Vanden-Eijnden, *J. Chem. Phys.* **125**, 084110 (2006).