JOHANNES GUTENBERG UNIVERSITY MAINZ

BACHELOR THESIS IN PHYSICS

Reweighting Dynamics out of Equilibrium for Steady State Systems

Author:
Timon WITTENSTEIN

Supervisor:
Dr. Tristan BEREAU

Examiners:
Prof. Dr. Kurt KREMER
Prof. Dr. Thomas SPECK

A thesis submitted to the Department of Physics, Mathematics and Computer Science (FB 08) of the Johannes Gutenberg University Mainz,

written in the

Max Plank Institute of Polymer Research
Biomolecular Simulations Group

July 9, 2018

Declaration of Authorship

I, Timon WITTENSTEIN, declare that this thesis titled, "Reweighting Dynamics out of Equilibrium for Steady State Systems" and the work presented in it are my own. I confirm that:

- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

Signed:

Date:

JOHANNES GUTENBERG UNIVERSITY MAINZ

Abstract

Max Plank Institute of Polymer Research
Biomolecular Simulations Group

Bachelor of Science in Physics

Reweighting Dynamics out of Equilibrium for Steady State Systems

by Timon WITTENSTEIN
twittens@students.uni-mainz.de

We present a maximum Caliber approach for reweighting transition rates of Markov state models for different non equilibrium steady state systems. In order to do this, we look at a reference model and reweight it into a model for another non equilibrium steady state system, only characterized by its average flux. For testing this method a minimalistic two dimensional toy model is used. The dynamical properties in this system, in particular the entropy production rate and the first passage time distributions, are compared for models obtained by the reweighting method and by simulation. Furthermore we introduce the local entropy production which gives us a more detailed look at the dynamics than the entropy production rate. This yields a relation similarly to detailed balance for non equilibrium steady states. We show that the reweighted Markov state models are able to predict the dynamics of the new system, but do not reproduce its stationary distribution correctly.

Acknowledgements

To begin with, I would like to thank my supervisor Dr. Tristan Bereau for offering me such an interesting topic of investigation and for his always highly competent remarks and suggestions.

Very special thanks to Marius Bause, a doctoral student in physics, for his eagerness and great assistance in order to accomplish the writing of this thesis.

Furthermore, I would like to thank my fellow student Felix Schreiber. Not only did he offer helpful feedback for my thesis, but he also made the last three years of my studies a truly enjoyable experience.

Last but not the least, my family are due my deep gratitude for their continuous support throughout my studies and life.

Contents

List of Figures

Chapter 1

Introduction

Markov state models of conformational dynamics have significantly advanced our understanding of biomolecular structure and dynamics [1]. They represent a master equation framework, i.e. they are able to describe the dynamics of a system in question. The master equation formalism has been used in many scientific fields [2]. In the Markov state model approach, the dynamics are modeled as a kinetic network of transitions between discretized states. This analysis works well with large collections of simulation trajectories which enable the identification of relevant metastable states and efficient sampling of the rates of transitions between them [3]. Furthermore, it allows to estimate the long timescale behavior which is otherwise not accessible by simulations. A complete description of the system is based on the estimates of the discrete time transition probability $p_{ij}(\tau)$, i.e., the probability of transitioning between states i and j in some time interval τ [4].

While these models are well studied in the equilibrium scheme, we will describe a systematic approach to construct Markov state models for systems driven out of equilibrium with dynamics that break detailed balance. Here we will focus on non equilibrium steady states where weaker conditions still hold true [5].

A non equilibrium steady state is the thermodynamic case where particle or energy fluxes still exist, but all physical variables are time independent. The work to maintain the steady state is provided by an external reservoir. It is one of the most important physical concepts for understanding biological processes [6]. This is not to say that all these processes are generally in steady states. In many cases though a steady state is a reasonable approximation to a short window of time or a localized region of interest, when populations, concentrations and stationary distributions are unchanging in time.

Entropy maximization principles provide a framework for understanding dynamics of systems in equilibrium. An off-equilibrium extension is the principle of maximum Caliber, which is a path entropy maximization applied to trajectories of dynamical systems [7]. It can be used to interpret dynamical fluctuations in biology and on the nanoscale, in single molecule and few particle systems by inferring stochastic dynamics from limited data [8]. The fundamental idea of the maximum

Caliber is to collect information about micro-trajectories such that it connects macroscopic quantities to these micro-trajectories. However, in principle these micro-trajectories are computationally not accessible in a non discretized space since their number diverges [7]. The space of trajectories is reduced by approximating the trajectories using Markovian transitions between predefined microstates. By inducing macroscopic information in the construction of the Caliber, we receive microscopic constraints. This information can be used to construct an off-equilibrium Markov state model.

We will introduce a minimal simulation model designed to test fundamental behavior of non equilibrium steady states [9]. Dynamical quantities like the entropy production [10] and first passage time distributions [11] are used to get macroscopic information about these steady states. It will be shown, how microscopic relations found by maximizing the Caliber are used to predict outcome of simulations. This opens opportunities to correctly sample otherwise not accessible rare events.

In the course of this thesis we will first build the theoretical framework on the simulation model, Markov state models in general and the principle of maximum Caliber. Using these principles we will derive a method to predict Markov state models for different systems given a reference model and the macroscopic average flux, used to characterize different non equilibrium steady states. The dynamical properties of these reweighted models will be compared to Markov models build directly out of a simulation trajectory to access the quality of the reweighting method. This will be done by analyzing the entropy production rate, as well as the first passage time distributions between metastable sets.

Chapter 2

Theory

2.1 Simulation Model

We are interested in examining a system in a non equilibrium steady state. In order to do this, we simulate a single particle in a two dimensional potential with periodic boundary conditions coupled to an equilibrium heat bath. Furthermore, we let a constant external force interact with the particle to simulate a non equilibrium steady state. This is achieved by using overdamped Langevin dynamics. Hereinafter the theoretical framework of such a simulation will be described shortly.

2.1.1 Overdamped Langevin Dynamics

The Langevin equation is a stochastic differential equation which is used to simulate a single particle/molecule in a solvent (like water), where the particle has much larger mass compared to the solvent. Therefore, the solvent molecules move much faster. These fast degrees of freedom are not explicitly taken into account, but implicitly by making use of a random force. We are not interested in the motion of the solvent molecules but rather in the motion of large particles (slow degrees of freedom). This means that we assume that the solvent molecules interact with the larger particles via collisions which occur at a characteristic time τ_{coll}. The Langevin equation gives a description for times $t \gg \tau_{coll}$ and takes the form (for simplicity we will focus on the one dimensional form but the expression can be easily expanded to a system of n particles in multiple dimensions) [12]:

$$m\dot{v} = F - m\gamma v + f_R, \tag{2.1}$$

where m is the particle mass, v the velocity of the large particle of interest, F the deterministic force field acting on the particle, γ the friction coefficient and f_R the random force. We can split the deterministic force F in a part induced by the potential V and another one representing the external force f_{ext}:

$$F = -\frac{\partial V}{\partial x} + f_{ext}. \tag{2.2}$$

The solvent is effectively replaced by an effective friction, described by a constant γ, and a random force. It can be shown that in order to reproduce the thermodynamic behavior of the solvent the random force has to fulfill the following equations [13]:

$$\langle f_R(t) \rangle = 0 \tag{2.3}$$

$$\langle f_R(t) f_R(t') \rangle = 2m\gamma k_B T \delta(t - t'), \tag{2.4}$$

where t and t' indicate two given times, k_B is the Boltzmann constant and T the temperature of the heat bath. Equation 2.3 ensures that the random force is unbiased and the delta function in equation 2.4 indicates that it is uncorrelated in time.

In the overdamped limit, which describes the high friction limit where the inertial term is neglected (set to zero), the Langevin equation 2.1 becomes (by also setting $m\gamma \equiv 1$):

$$0 = -\frac{\partial V(x)}{\partial x} + f_{ext} - \dot{x} + f_R. \tag{2.5}$$

This equation is now describing Brownian dynamics [13]. We will use the overdamped limit in order to drive the system further out of equilibrium while still keeping track on the direction the particle is heading. Since we will use periodic boundary conditions, without the overdamped limit, we could run into problems here since the path of a transition not well defined anymore if we sample the system for higher times $\tau = n\Delta t$, where n is an integer factor of the integration time step Δt.

2.1.2 Numerical Integration

Following [12] we can derive an expression for a first order integrator for the overdamped Langevin equation 2.5 if we assume small discrete time steps Δt:

$$x(t + \Delta t) = x(t) + F(x(t))\Delta t + \sqrt{2k_B T}\zeta\sqrt{\Delta t}, \tag{2.6}$$

where ζ is given by a Gaussian distributed random number with zero mean and unit variance. In practice, this number needs to be generated at every time step Δt.

2.2 Markov State Models

Markov state models are used to describe the entire dynamics of a system, in our case assumed to be in a thermodynamic non equilibrium steady state. This model can be represented by a $n \times n$ transition probability matrix, where the entire configuration space spanned by the system has been divided into n states. It is often constructed using a trajectory tracking the dynamical progress of the system (for example molecular dynamics simulation) by checking which states the system occupies at time steps separated by the so called *lag time* τ. For the transition probability

matrix to be Markovian, the system must be "memoryless". This means the probability to jump from a state i to a state j does not depend on its history prior to state i [14].

2.2.1 Markov Model Theory

2.2.1.1 The Transfer Operator

Consider a state space Ω containing all variables needed to describe the current state of the system. We will first treat the more general case of a continuous state space. $\mathbf{x}(t) \in \Omega$ denotes the state of the system at time t. The dynamical process $(\mathbf{x}(t))_{t\in T}$, $T \subset \mathbb{R}_{0+}$ is therefore continuous in space and can be either time continuous or time discrete (for computational purposes this will be the case). We will assume that the dynamic process has the following properties [4]:

1. $(\mathbf{x}(t))_{t\in T}$ is a Markov process in the full state space Ω. This means that $\mathbf{x}(t+\Delta t)$ is calculated based on $\mathbf{x}(t)$ alone and does not require the previous history. In other words: A Markov process is a random process in which the future is independent of the past, given the present. Thus, Markov processes are the natural stochastic analogs of processes described by equations of motion. [15].

   In addition, we assume the transition probability density $p(\mathbf{x}, \mathbf{y}; \tau)$ is well defined. $p(\mathbf{x}, \mathbf{y}; \tau)$ tells us the probability of the trajectory to go from state $\mathbf{x}$ to $\mathbf{y}$ in time τ, or more precisely, the density associated with it so that

   $$\int_\Omega \mathrm{d}\mathbf{y}\; p(\mathbf{x}, \mathbf{y}; \tau) = 1 \tag{2.7}$$

   holds true. Such a probability density for a one dimensional diffusion process is depicted in figure 2.1(b).

2. $\mathbf{x}(t)$ is ergodic, i.e. the space Ω does not have two or more subsets that are dynamically disconnected and for $t \to \infty$ each state $\mathbf{x}$ will be visited infinitely often. The process therefore has a normalized stationary density $\mu(\mathbf{x}) : \Omega \to \mathbb{R}_+$ [see figure 2.1(a)] such that

   $$\mu(\mathbf{y}) = \int_\Omega \mathrm{d}\mathbf{x}\; p(\mathbf{x}, \mathbf{y}; \tau)\, \mu(\mathbf{x}). \tag{2.8}$$

These conditions are rather general and certainly hold true for the overdamped Langevin dynamics in both the equilibrium and the non equilibrium steady state case.

Now we look at an ensemble, distributed by a probability density $p_t(\mathbf{x})$ in the state space Ω, instead of a single trajectory. After some time τ the system undergoes transitions in state space according to the transition probability density $p(\mathbf{x}, \mathbf{y}; \tau)$. This change of the probability density $p_t(\mathbf{x})$ to $p_{t+\tau}(\mathbf{x})$ can be described as a continuous transfer operator which we can define as the following:

$$u_{t+\tau}(\mathbf{y}) = \hat{T}(\tau) \circ u_t(\mathbf{y}) = \frac{1}{\mu(\mathbf{y})} \int_\Omega \mathrm{d}\mathbf{x}\, p(\mathbf{x}, \mathbf{y}; \tau)\, \mu(\mathbf{x})\, u_t(\mathbf{x}). \tag{2.9}$$

$\hat{T}(\tau)$ does not propagate probability densities but instead functions $u_t(\mathbf{x})$ that differ by a factor of the stationary density $\mu(\mathbf{x})$ [4]:

$$p_t(\mathbf{x}) = \mu(\mathbf{x})\, u_t(\mathbf{x}). \tag{2.10}$$

This transfer operator has the following properties:

- $\hat{T}(\tau)$ can be used to propagate the dynamics to arbitrarily long times $t + k\tau$ since it fulfills the Chapman-Kolmogorov equation: [4]

$$u_{t+k\tau}(\mathbf{x}) = \left[\hat{T}(\tau)\right]^k \circ u_t(\mathbf{x}). \tag{2.11}$$

- We consider the transfer operator in the Hilbert space of square integrable, μ-weighted functions L^2_μ with the scalar product:

$$\langle u, v \rangle_\mu = \int_\Omega \mathrm{d}\mathbf{x}\, u^*(\mathbf{x})\, v(\mathbf{x})\, \mu(\mathbf{x}). \tag{2.12}$$

  In this space $\hat{T}(\tau)$ has eigenfunctions $\Psi_i(\mathbf{x})$ with the associated eigenvalues λ_i:

$$\hat{T}(\tau) \circ \Psi_i(\mathbf{x}) = \lambda_i \Psi_i(\mathbf{x}). \tag{2.13}$$

  The length of these eigenfunctions is defined by the normalization condition $\langle \Psi_i, \Psi_i \rangle_\mu = 1$. We can now introduce the left eigenfunctions $\Phi_i(\mathbf{x})$ which correspond to the probability density again and as such are related to $\Psi_i(\mathbf{x})$ by a factor of the stationary density $\mu(\mathbf{x})$:

$$\Phi_i(\mathbf{x}) = \mu(\mathbf{x})\, \Psi_i(\mathbf{x}). \tag{2.14}$$

  In the case of a sufficiently fine discrete state space we can approximate $\hat{T}(\tau)$ by a transition matrix $\mathbf{T}(\tau)$ (more on that in section 2.2.1.2). $\Phi_i(\mathbf{x})$ and $\Psi_i(\mathbf{x})$ are then approximated by the left and right eigenvectors of that transition matrix respectively. This approximated transition matrix is per definition a stochastic matrix [16] and has as such only eigenvalues (real or complex) with $|\lambda_i| \leq 1$. Furthermore, 1 is always an eigenvalue of this matrix [17].

  Using such a state space with a suitable lag time τ (more on this later in section 2.2.2.2) the largest of these eigenvalues correspond to the largest ones of

the actual transfer operator $\lambda_1 = 1 \geq |\lambda_2| \geq |\lambda_3| \geq ... \geq |\lambda_m|$. This finite number of m isolated, *dominant* eigenvalue/eigenfunction pairs form the so called *discrete* spectrum of the transfer operator. All the eigenvalues with $|\lambda_i| \leq |\lambda_m|$ (infinitely many ones) form the *essential* spectrum, which is the fast decaying part of the transfer operator, and as such correspond to all fast processes which are usually not of interest [4].

- The eigenfunction corresponding to the largest eigenvalue $\lambda = 1$ is a constant function on all state space Ω [see figure 2.1(c), top] [17]:

$$\hat{T}(\tau) \circ \mathbf{1} = \mathbf{1} = \Psi_1(\mathbf{x}) \tag{2.15}$$

and due to relation 2.14 $\Phi_1(\mathbf{x})$ corresponds to the stationary density $\mu(\mathbf{x})$ [see figure 2.1(d), top]

If the Markov process is reversible (equilibrium case) the *detailed balance* condition

$$\mu(\mathbf{x})\, p(\mathbf{x}, \mathbf{y}; \tau) = \mu(\mathbf{y})\, p(\mathbf{y}, \mathbf{x}; \tau) \tag{2.16}$$

is satisfied and the transfer operator is self adjoint. In this case all the eigenvalues λ_i are real valued. Also different eigenfunctions are orthogonal [18]:

$$\left\langle \Psi_i, \Psi_j \right\rangle_\mu = \delta_{ij}. \tag{2.17}$$

To see the significance of the eigenvalue/eigenfunction pairs in the reversible case we separate the dynamics of the system into a superposition of the discrete and essential spectrum. The discrete spectrum can, again, be decomposed into a superposition of m individual slow dynamical processes. This yields:

$$u_{t+k\tau}(\mathbf{x}) = \hat{T}_{slow}(k\tau) \circ u_t(\mathbf{x}) + \hat{T}_{fast}(k\tau) \circ u_t(\mathbf{x}) \tag{2.18}$$

$$= \sum_{i=1}^{m} c_i \lambda_i^k \Psi_i(\mathbf{x}) + \hat{T}_{fast}(k\tau) \circ u_t(\mathbf{x}), \tag{2.19}$$

where c_i are the Fourier coefficients of this superposition [4]. [1]

In this equation we can associate the eigenfunctions $\Psi_i(\mathbf{x})$ of the discrete spectrum to different dynamical processes which decay with increasing time index k. In the long time limit $k \to \infty$ only the first term with $\lambda_1 = 1$ remains and the system relaxes to its stationary density. Thus we can identify a physical timescale for the eigenvalues with $\lambda \neq 1$, indicating how quickly the process is decaying (the closer λ_i is to 1 the slower it decays) [see figure 2.1(e)]. If we define the timescales as the

[1] It shall be noted that this decomposition is only valid in the equilibrium case. In the Markov state models for the non equilibrium steady states the detailed balance condition 2.16 is not fulfilled and the eigenfunctions are not orthogonal anymore. Furthermore, the eigenvalues not equal to 1 can become complex. This leaves their physical interpretation still open for discussion.

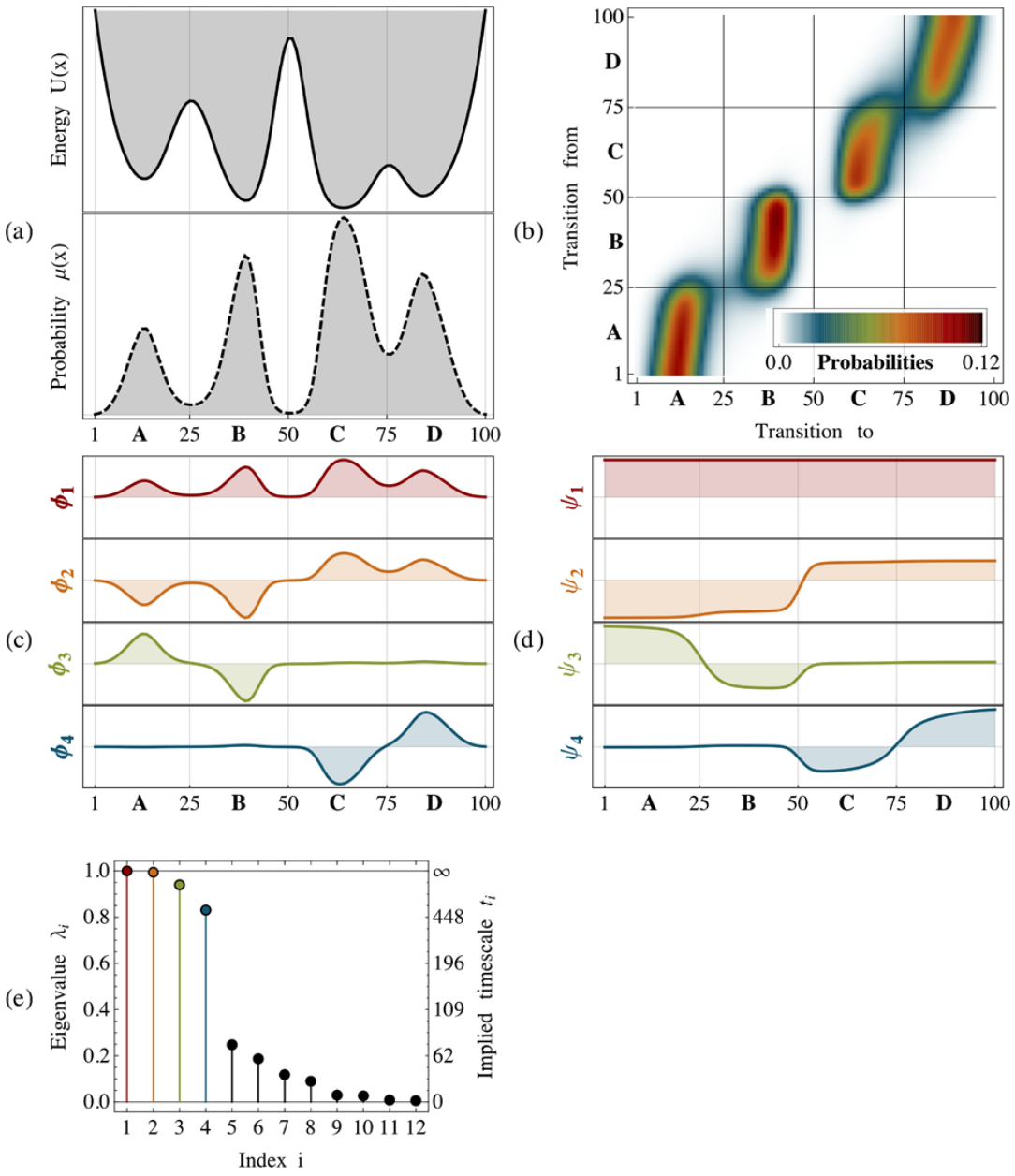

FIGURE 2.1: (a) A potential energy function with four metastable sets A, B, C and D for a diffusion-in-potential dynamics and the corresponding stationary density. (b) Density Plot of the transfer operator. One can clearly see the diagonal block structure where the transition density is large within the four metastable sets and very low in between those. (c) The four dominant eigenfunctions of the transfer operator weighted with the stationary density. (d) The four dominant eigenfunctions associated with different dynamical processes at different timescales. Ψ_2 for example indicates the slowest transition between the sets A+B and C+D. (e) Eigenvalues of the transfer operator. A gap between the slow ($\lambda_i \approx 1$) and the fast processes is visible. Reprinted from JH Prinz et al. [4], with the permission of AIP Publishing.

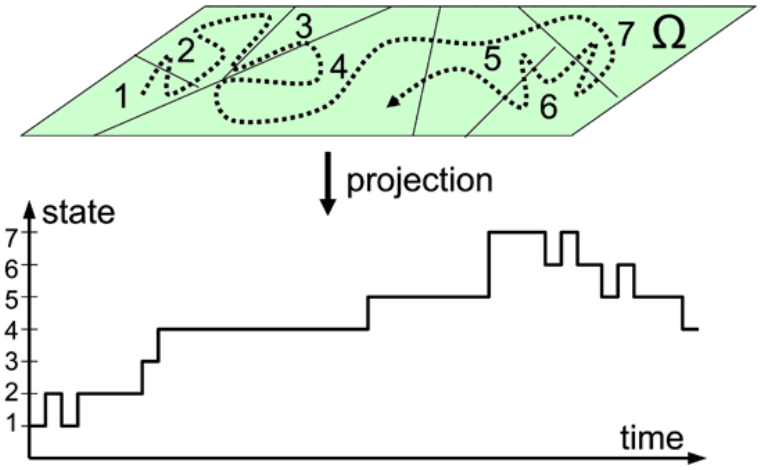

FIGURE 2.2: The true dynamics (dashed line) is projected onto a discrete state space. The resulting Markov State Model approximates this as a Markov jump process. Reprinted from JH Prinz et al. [4], with the permission of AIP Publishing.

following:

$$t_i = -\frac{\tau}{\ln \lambda_i}, \tag{2.20}$$

we can rewrite equation 2.19 in terms of these implied timescales:

$$u_{t+k\tau}(\mathbf{x}) = 1 + \sum_{i=2}^{m} c_i \exp\left(-\frac{k\tau}{t_i}\right) \Psi_i(\mathbf{x}) + \hat{T}_{fast}(k\tau) \circ u_t(\mathbf{x}). \tag{2.21}$$

In this equation t_i inherits the role of a timescale, thus, legitimating our definition in equation 2.20. This implies that there are simultaneous dynamical processes acting on different timescales if there are gaps amongst the first m eigenvalues. For example if $\lambda_1 = 1$, $\lambda_2 \sim 1$ and $\lambda_3 \ll 1$ the system has a two state kinetics [4].

2.2.1.2 Discretization of State Space

Markov State Models are appropriate discretizations of the eigenvalue problem that allow to approximate the dominant eigenvalues and eigenvectors of the transfer operator $\hat{T}(k\tau)$, and thus, describe the kinetics of long timescale processes using an underlying Markov process [18]. This is achieved using the Galerkin approach [19]:

Looking at a Markovian jump process between a set of states, we consider a discretization of state space Ω into k crisp sets [2] (figure 2.2). This means that the probability of jumping to a new state only depends on the current state and, as such, the dynamics are Markovian and can be completely described by a transition matrix $\mathbf{T}$ [20]:

$$T_{ij} = P\left(\mathbf{x}_{t+\tau} \in s_j | \mathbf{x}_t \in s_i\right). \tag{2.22}$$

In this equation the states $\{s_i\}_{i=1}^{k}$ are a set of non overlapping subsets of Ω that partition the space as follows: $s_i \subseteq \Omega \; \forall \; i$, $\cup_{i=1}^{k} s_i = \Omega$ and $s_i \cap s_j = \emptyset \; \forall \; i \neq j$.

[2] In some cases it can be desirable to use a fuzzy partitioning but that case is not important for this Thesis.

Using a partitioning function:

$$\chi_i(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in s_i \\ 0 & \mathbf{x} \notin s_i \end{cases} \tag{2.23}$$

we can get a relation for the transition matrix $\mathbf{T}(\tau) \in \mathbb{R}^{k \times k}$ [20]:

$$T_{ij} = \frac{\int_{s_j} \mathrm{d}\mathbf{x}\, \mu(\mathbf{x}) \left(\hat{T}(\tau) \circ \chi_i(\mathbf{x})\right)}{\int_{s_i} \mathrm{d}\mathbf{x}\, \mu(\mathbf{x})}. \tag{2.24}$$

Each element of T_{ij} describes the probability to find the system in state j at the time $t + \tau$ given it was in state i at time t. The stationary probability π_i to be in state i is given by:

$$\pi_i = \int_{s_i} \mathrm{d}\mathbf{x}\, \mu(\mathbf{x}) \tag{2.25}$$

or respectively, as the i-th entry of the first left eigenvector of the transition matrix [4].

Modeling the original process with this discrete state space Markov process is an approximation and, as such, involves a systematic *discretization error*. In order to keep it small, a suitable lag time has to be determined (more on this in section 2.2.2.2).

We can now relate the functions that get propagated by the transfer operator $\hat{T}$ (u_t in equation 2.9) to the right eigenvectors of the matrix Tij and the probability densities p_t (equation 2.10) as the left eigenvectors. This means that if we consider a probability vector to be in state j at time t: $\mathbf{p}(t) \in \mathbb{R}^k$, after time τ the distribution of the probabilities will develop according to [4]:

$$p_j(t + \tau) = \sum_{i=1}^{k} p_i(t)\, T_{ij}(\tau), \tag{2.26}$$

or in matrix form:

$$\mathbf{p}^T(t + \tau) = \mathbf{p}^T(t)\, \mathbf{T}(\tau). \tag{2.27}$$

2.2.2 Building a Markov State Model

In this section we will discuss the problem of estimating the transition matrix $\mathbf{T}(\tau)$, since in most practical cases it is not obtained by directly discretizing the transfer operator. Instead, we can build a Markov Model by using a set of trajectory data, a discretization of state space Ω and selecting an appropriate lag time τ (our observation interval) [21].

2.2.2.1 Transition Count Matrix

Consider a trajectory X with N configurations sampled at a fixed time Δt:

$$X = \mathbf{x}_1 = \mathbf{x}\left(t = 0\right), \mathbf{x}_2 = \mathbf{x}\left(t = \Delta t\right), ..., \mathbf{x}_N = \mathbf{x}\left(t = (N-1)\Delta t\right). \tag{2.28}$$

Using a state space discretization $[s_1, ..., s_K]$ with K microstates, we can assign each element of the trajectory to a discrete state s_i. All the trajectory information can now be stored as a sequence of discrete states. Constructing a count matrix with this data is now simply a matter of counting the transitions between two states at times $\tau = l\Delta t$ where $l \in \mathbb{N}$. In doing so, we have to distinguish between two different counting schemes (figure 2.3): Either the trajectory is sampled at lag time τ and only these sample points are used for counting or a count window of length τ is shifted along the time line. On the on hand most of the data is ignored in the first approach which can lead to numerical problems. On the other nearby transitions in the second one are not statistically independent and the resulting count Matrix is therefore harder to deal with later on. In our case we have sufficient sampling to use the first method without any problems [4].

We can now define a state to state count matrix $\mathbf{C}\left(\tau\right) = \left[c_{ij}\left(\tau\right)\right]$ at lag time τ. Each element is given by:

$$c_{ij}\left(\tau\right) = c_{ij}\left(l\Delta t\right) = \sum_{k=0}^{[(N-1)/l]-1} \chi_i\left(\mathbf{x}_{lk+1}\right) \chi_i\left(\mathbf{x}_{l(k+1)+1}\right). \tag{2.29}$$

Therefore, each element of this count matrix describes the number of times a transition from state i to state j was observed. If we have multiple trajectories instead of one we can simply add the resulting count matrices [4].

Intuitively we can estimate a true stochastic transition matrix by:

$$T_{ij}\left(\tau\right) = \frac{c_{ij}\left(\tau\right)}{c_i\left(\tau\right)}. \tag{2.30}$$

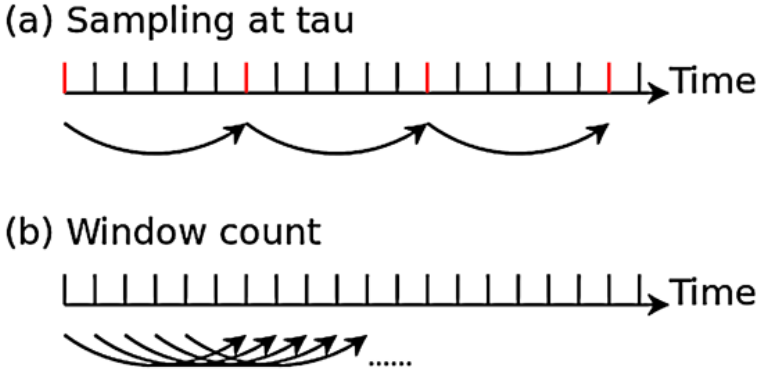

FIGURE 2.3: Transition counting schemes for estimating the count Matrix. Reprinted with permission from GR Bowman et al. [21], Chapter 4.

In this equation $c_i(\tau)$ describes the row sums of $\mathbf{C}(\tau)$ which is simply the number of times the trajectory (sampled at τ) visited state i:

$$c_i(\tau) = \sum_{j=1}^{K} c_{ij}(\tau). \tag{2.31}$$

However, equation 2.30 is only true for infinite sampling. In practice estimating transition matrices is more complicated due to a number of issues, like finite sampling and imperfections in microstate definitions. For example in the equilibrium case detailed balance ($\pi_i T_{ij} = \pi_j T_{ji}$ in terms of the discrete description) is not necessarily fulfilled even though it is a physical property of the system. Therefore, it has to be enforced on the transition matrix (by Bayesian model selection approaches[3]).

Nonetheless, we will stick to equation 2.30 in this thesis in order to determine the transition matrix. Our system is simple and the sampling rate high enough to keep the error obtained from this estimation neglectable. Furthermore, the detailed balance condition which yield the most information is not fulfilled in non equilibrium steady states.

2.2.2.2 Validation of the Markov State Model

In this section we will discuss how to choose an appropriate lag time τ for our Markov State Model.

It is important that the assumption we made when discretizing the state space, that the underlying process is still a Markov process, holds true. So in order to test the validity of the model we can use the Chapman Kolmogorov equation 2.11. It can be reformulated in terms of our transition matrix $\mathbf{T}(\tau)$:

$$\mathbf{T}(n\tau) = \mathbf{T}(\tau)^n, \tag{2.32}$$

where n is an integer number of steps.

Plotting the slowest relaxation timescale (equation 2.20) of a model as a function of the lag time, we can use the Chapman Kolmogorov equation to provide some model validation and an estimated lag time. Based on the Markov assumption, the relaxation times of a Markov model with lag time $n\tau$ should be the same as the ones with lag time τ [21]:

$$t_i = \frac{n\tau}{\ln \lambda_{i,\mathbf{T}(n\tau)}} = \frac{n\tau}{\ln \lambda^n_{i,\mathbf{T}(\tau)}} = \frac{\tau}{\ln \lambda_{i,\mathbf{T}(\tau)}}. \tag{2.33}$$

In the second step we made use of the fact that the Chapman Kolmogorov equation is also true for the eigenvalues corresponding to the transition matrices (multiply equation 2.32 with each eigenvector Ψ_i).

[3] Further information on how to do this can be found in Bowman [21], chapter 4.

So beyond our chosen lag time τ the timescale plot should converge. This lag time should be chosen as small as possible to achieve the finest time resolution possible for our model.

Note that while Markovian dynamics (Chapman Kolmogorov equation 2.32 holds true) implies that the timescales will be constant in τ, the reverse is not necessarily true. However, constant implied timescales are a very strong indicator that the Markov state model approximates the underlying dynamics well [22]. To provide full validation of the model one also needs to conduct the Chapman-Kolmogorov test as described in JH Prinz et al. [4].

In this test we are interested in checking whether the approximation

$$[\mathbf{T}(\tau)]^k \approx \mathbf{T}(k\tau), \tag{2.34}$$

holds within statistical uncertainty.

At this point the importance of a proper state space discretization becomes apparent[4] in order to keep the already mentioned discretization error as small as possible. If the error is large the Markov assumption is violated and the error has to be compensated by a larger lag time.

2.3 Maximum Caliber

2.3.1 Maximum Entropy for Dynamical Pathways

The principle of maximum Caliber is an extension of the principle of maximum entropy and not limited to states of equilibrium. Rather than seeking probability distributions of static observables (like energy, magnetization...), we now want to find probability distributions of dynamical observables. This is done by maximizing a path entropy over all possible dynamical pathways which are subject to an observable quantity like an average velocity or a flux.

For problems of this type we can define the path entropy as [7]:

$$H(\{p_C\}) = -\sum_C p_C \log p_C, \tag{2.35}$$

where p_C is the probability for a certain dynamical pathway following path C. Furthermore, we suppose constraints on the dynamics:

$$F^{(\alpha)}(p_C) = 0, \tag{2.36}$$

where α is indexing the number of constraints. These constraints also include normalizations. In order to analytically handle the maximum Caliber, we are mostly

[4]There are a number of Methods to construct a kinetically relevant clustering using geometric criteria(Bowman [21],chapter 2) and there is not necessarily a single right answer. However, these are not relevant in the course of this thesis since we only use a fairly simple clustering.

interested in constraints that are linear in p_C:

$$\sum_C A_C^{(\alpha)} p_C - \overline{A}(\alpha) = 0. \tag{2.37}$$

$\overline{A}(\alpha)$ describes the measured average of the quantity $A_C^{(\alpha)}$ over the paths C.

The principle of maximum Caliber is now to maximize the entropy over pathways (equation 2.35), given the constraints introduced in equation 2.37. Admittedly the principle is formalistic and usually not explicitly computable for continuous pathways C since the number of trajectories goes to infinity. However, if a set of pathways is specified, like in the case of discrete states, maximum Caliber can be readily applied [7].

2.3.1.1 Maximization of Path Entropy for discrete Paths

We now consider trajectories which are composed of discrete time steps and also mapped to discrete states. In this case a particular pathway is described by a sequence $C = \{i_0, i_1, ..., i_T\}$ with a total number of time steps T. i_x describes the state occupied by the system at time step x. If we now assume a Markov process the probability for a discrete path is given by [7]:

$$p_C = p_{i_0}\ p_{i_0 i_1}\ p_{i_1 i_2} \cdots p_{i_{T-1} i_T}, \tag{2.38}$$

where p_{ij} is the transition probability from state i to state j and p_i a single state occupation probability. These transition probabilities correspond to the transition matrix of a Markov State model described in the previous section. The single state occupation probability corresponds to i-th entry of the stationary distribution π_i.

If we plug this probability in equation 2.35 and express the logarithm as a sum over each individual terms we obtain:

$$H(T) = -\sum_{\{i_0,i_1,...,i_T\}} \pi_{i_0} p_{i_0 i_1} ... p_{i_{T-1} i_T} \log p_{i_0} - \sum_{\{i_0,i_1,...,i_T\}} \pi_{i_0} p_{i_0 i_1} ... p_{i_{T-1} i_T} \log p_{i_0 i_1} - ... \,. \tag{2.39}$$

This equation can be simplified assuming time independent distributions and transition probabilities as it is the case for an equilibrium system as well as a non equilibrium steady state system. We do so by using a normalization and the so called *global balance* condition [7], characteristic for a non equilibrium steady state:

$$\sum_j p_{ij} = 1, \tag{2.40}$$

$$\sum_i \pi_i p_{ij} = \pi_j. \tag{2.41}$$

Performing the summation over all indices except i_0 in the first term of equation 2.39 using these conditions yields $-\sum_i \pi_i \log p_i$ with a simplified notation for the

dummy indices. In the second term we sum over all indices expect i_0 and i_1. All the other terms can be treated similarly and each term reduces to $-\sum_{ij} \pi_i p_{ij} \log p_{ij}$. All together we obtain:

$$H(T) = -\sum_i \pi_i \log p_i - T \sum_{ij} \pi_i p_{ij} \log p_{ij}. \tag{2.42}$$

For large T the first term is negligible and we can approximate the path entropy per step as [7]:

$$H = -\sum_{ij} \pi_i p_{ij} \log p_{ij}. \tag{2.43}$$

Here we derived the form of the path theory assuming Markov chain kinetics. By maximizing the Caliber given all constraints of the system we would obtain our Markov State model characterized by p_{ij}.

2.3.2 Reweighting Dynamics

In theory the maximum Caliber as described above allows us to build a Markov State model using constraints via Lagrangian multipliers. However, this requires full knowledge of which constraints are dominant for a given system. This is usually not possible for more complex systems. Instead,we can make use of the *relative path entropy*.

2.3.2.1 Relative Path Entropy

Suppose a Markov State model defined by the transition probabilities q_{ij}, the *prior*, is perturbed. Our goal is to find a new model with new transition probabilities p_{ij} and a new stationary distribution π_i. The Lagrangian multipliers now represent how much the original system is altered by the given constraints which describe the perturbation.

In general the information on a distribution p with respect to another distribution q is quantified by the relative entropy, or Kullback-Leibler divergence [3]:

$$D(p|q) = \sum_x p(x) \log \frac{p(x)}{q(x)}. \tag{2.44}$$

For a Markov chain the relative path entropy is then given by:[5]

$$D = \sum_{i,j} \pi_i p_{ij} \log \frac{p_{ij}}{q_{ij}}. \tag{2.45}$$

Minimizing this relative path entropy of the perturbed system with respect to the unperturbed one is equivalent to maximizing the Caliber [3].

[5]The derivation of this formula is in conjunction to the one of the path entropy for Markov chains in the previous section.

As an example consider the simulation of a system in equilibrium, and thus, with no external force and no average flux in order to use this information to build a reference model q_{ij}. Minimizing the relative path entropy with constraints that characterize a system out of equilibrium, like introducing a time-independent average flux describing a non equilibrium steady state, allows us to predict the Markov state model of this new system.

Minimizing the relative path entropy forces the new Markov state model to be close to the reference model. Considering that there exist multiple Markov state models fulfilling the given constraints, this is a reasonable assumption. The prior q_{ij} contains dynamical information about the system we want to preserve.

In the following we want to derive an analytic expression of the new transition matrix.

2.3.2.2 Macroscopic average Flux

In the course of this thesis we are interested in reweighing different non equilibrium steady state systems (including equilibrium) into each other. We can introduce the macroscopic average flux $\langle J \rangle$ as an experimentally measurable quantity. This average flux is sufficient to describe the perturbation on our system (later on we will validate this assumption), and thus, can quantify the distance of a system to thermal equilibrium. Therefore a microscopic description of the average flux is needed:

$$\langle J \rangle = \frac{1}{\tau} \sum_{i,j} \pi_i p_{ij} F_{ij}. \tag{2.46}$$

In this equation F_{ij} describes the distance from state i to state j projected onto the direction of the external force. The sign gets determined by the direction of the transition in question. In the direction of the force it will be positive and otherwise negative. Since we deal with a system containing periodic boundary conditions this direction is not necessarily clear. However, our lag time is small enough that it is unlikely (in practice zero) for a single transition to cover the whole system (or rather half of it). Therefore the direction of the transition is well defined. F_{ij} also introduces a distance metric on the macroscopic average flux.

The prefactor $1/\tau$, on the other hand, introduces a timescale in the description of the macroscopic average flux. This allows the comparison of this macroscopic average flux to, for example, the average flux calculated from a simulated trajectory or an average flux obtained from experimental results.

2.3.2.3 Derivation of the Reweighed Markov State Model

This derivation follows the approach suggested in [23] applied to our problem.

In addition to the constraint describing the perturbation to the reference system, we also need normalization constraints as well as a constraint enforcing the global balance condition 2.41.[6]

All combined this yields a Caliber of the following form:

$$C = \underbrace{-\sum_{ij} \pi_i p_{ij} \log \frac{p_{ij}}{q_{ij}}}_{\text{Relative Path Entropy}} + \underbrace{\sum_i m_i \left(\sum_j \pi_i p_{ij} - \pi_i \right) + \zeta \left(\sum_{ij} \pi_i p_{ij} - 1 \right)}_{\text{Normalization}} + \underbrace{\sum_j n_j \left(\sum_i \pi_i p_{ij} - \pi_j \right)}_{\text{Global Balance}} - \underbrace{\gamma \left(\sum_{ij} \pi_i p_{ij} F_{ij} - \langle J \rangle \right)}_{\text{External Constraints (e.g. Flux)}} . \quad (2.47)$$

We will now maximize this with respect to p_{ij} and π_i: Differentiating the Caliber with respect to p_{ij} yields the following relation:

$$\frac{\partial C}{\partial p_{ij}} = \sum_{i,j} \pi_i \left(-\log \frac{p_{ij}}{q_{ij}} - 1 + m_i + n_j + \zeta - \gamma F_{ij} \right) \stackrel{!}{=} 0 \quad (2.48)$$

$$\Rightarrow p_{ij} = q_{ij}\, e^{m_i + n_j + \zeta - 1 - \gamma F_{ij}}. \quad (2.49)$$

On the other hand, differentiating the Caliber with respect to π_i, we get:

$$0 = -\sum_j p_{ij} \log p_{ij} + \sum_j p_{ij} m_i - m_i + \sum_j p_{ij} n_j - n_i + \zeta \sum_j p_{ij} - \gamma \sum_j p_{ij} F_{ij}. \quad (2.50)$$

We can now substitute the result we obtained in equation 2.49. Together with the normalization condition $\sum_j p_{ij} = 1$, we get another relation, this time for the Lagrangian multipliers:

$$m_i + n_i = 1. \quad (2.51)$$

At this point we introduce a matrix $W_{ij} = q_{ij}\, e^{-\gamma F_{ij}}$, the vectors $\Phi_i = e^{-m_i}$ and respectively, using the relation above, $\Phi_j = e^{-m_j} = e^{n_j - 1}$ as well as a scalar $\eta = e^{-\zeta}$. Using this notation we get a reformulation for the maximum path entropy transition probabilities:

$$p_{ij} = \frac{\Phi_j}{\eta \Phi_i} W_{ij}. \quad (2.52)$$

If we now impose the normalization constraint $\sum_j p_{ij} = 1$ once again, the reason

[6] At this point one could ask why we did not enforce these constraints on the transition matrix we obtained in section 2.2.2.1. By our definition 2.30 the transition matrix is already normalized. On the other hand the global balance condition, which is an intrinsic property of a steady state, is not necessarily fulfilled. Though by testing this condition with the model one realizes that it is only violated by a factor of around $\sim 10^{-10}$ and as such can be ignored.

for this reformulation becomes apparent: Our problem gets reduced to an eigenvalue problem only dependent on one Lagrangian multiplier γ and q_{ij}:

$$1 = \sum_j \frac{\Phi_j}{\eta \Phi_i} W_{ij}(\gamma) \;\Leftrightarrow\; \sum_j W_{ij}(\gamma)\,\Phi_j = \eta \Phi_i \tag{2.53}$$

Since W_{ij} is per definition a positive matrix, we can apply the Perron-Frobenius theorem [17] which guarantees that the maximum eigenvalue of $\mathbf{W}$ is positive and the corresponding eigenvector has only positive elements. We now choose η to be the maximum eigenvalue and $\vec{\Phi}$ the corresponding eigenvector. This way the transition probabilities are guaranteed to be positive (negative transition probabilities have no physical interpretation).

By defining the desired system with its average flux $\langle J \rangle$, we find a corresponding Lagrangian multiplier γ that reweights the system accordingly. Thus, with equation 2.52, we have an analytic expression for our new Markov state model characterized by p_{ij}.

2.4 Dynamical properties

The intention of this thesis is to test the quality of the reweighting method described in the previous section. In order to do this, we compare the dynamical properties of the reweighted Markov state model with a model obtained by a simulation trajectory. These properties include the first passage time and the entropy production rate of each model.

2.4.1 First Passage Time Distribution

First passage times are widely used to characterize different stochastic processes like chemical reactions, protein folding and diffusion processes since it contains key information on the kinetics of any process. It gives an approximate description (in the form of a probability distribution) on the timescale of a certain process. [11]

In the context of a Markov state model, the first passage time is a probability distribution associated with the discrete state space. It can be derived from the transition matrix p_{ij}:

$f_{ij}^{(n)}$ will now refer to the probability that the first passage time to get from state i to state j is equal to $n\tau$. Here τ, again, denotes the lag time of our model. By definition, we get that $f_{ij}^{(1)} = p_{ij}$, since p_{ij} already describes the probability of the transition $i \to j$ in the time τ. To get that value for $f_{ij}^{(2)}$, we need to sum over all possible paths over a third state $m \neq j$ (otherwise n would be 1). Thus, we can derive it from $f_{ij}^{(1)}$ and the probability p_{im} [11]:

$$f_{ij}^{(2)} = \sum_{m \neq j} p_{im} f_{mj}^{(1)}. \tag{2.54}$$

Similarly, we get relations for any $f_{ij}^{(n)}$ from $f_{ij}^{(n-1)}$ for all $n > 1$ following a recursive formula:

$$f_{ij}^{(n)} = \begin{cases} p_{ij} & for \;\; n = 1 \\ \sum_{m \neq j} p_{im} f_{mj}^{(n-1)} & for \;\; n > 1 \end{cases} \tag{2.55}$$

This formula yields a first passage time distribution in time steps of τ. In order to compare different models, it is suitable to also have a look at the moments of this distribution. Hereby the first moment corresponds to the mean first passage time of the transition in question.

2.4.2 Entropy Production

Driving a system away from thermal equilibrium implies a non-vanishing entropy production which therefore describes a macroscopic property of a non equilibrium steady state. Its importance for the system dynamics and its appliance to Markov state models will be discussed hereinafter.

The Entropy production in the system arises from non-vanishing probability currents $\Phi_{ij} - \Phi_{ji}$ where $\Phi_{ij} = \pi_i p_{ij}$ represents probability fluxes along a transition from state i to j. Following J. Schnakenberg [10] the total entropy production rate is given by:

$$\dot{S}_{tot} = \sum_{ij} \Phi_{ij} \log \frac{\Phi_{ij}}{\Phi_{ji}} = \underbrace{\sum_{ij} \Phi_{ij} \log \frac{\pi_i}{\pi_j}}_{\dot{S}_{sys}} + \underbrace{\sum_{ij} \Phi_{ij} \log \frac{p_{ij}}{p_{ji}}}_{\dot{S}_{med}} \tag{2.56}$$

In thermal equilibrium, the condition of detailed balance $\Phi_{ij} = \Phi_{ji}$ holds true which implies zero entropy production.

The first term describes the time derivative of Gibbs Entropy production $S_{sys} = -\sum_i \pi_i \log \pi_i$ and thus represents the entropy production caused by the system itself. For non equilibrium steady states this term vanishes since the stationary distribution is time independent. Consequently $\dot{S}_{tot} = \dot{S}_{med}$ where the second term represents the coupling of the system with the medium. The coupling can be, for example, caused by a heat or particle flux flowing from the medium into the system [24].

The entropy production describes the asymmetry in the probability fluxes and thus is an important property of a non equilibrium steady state and has to be reproduced by a reweighted Markov state model.

2.4.2.1 Local Entropy Production

To get more insight on the microscopic level inside the system we want to take a closer look at an entropy production between two states. We will call this the local entropy production. This entropy production Δs_{ij} of the medium between two

discrete states i and j is given by the logarithm of the ratio of probabilities for the forward and backward path, p_{ij} and p_{ji} respectively [9, 25]:

$$\Delta s_{ij} = \log \frac{p_{ij}(\tau)}{p_{ji}(\tau)}. \tag{2.57}$$

By assuming a single trajectory going from i to j this ratio can be calculated in terms of the path integral representation of the Langevin equation [26]:

$$\log \frac{p_{ij}(\tau)}{p_{ji}(\tau)} = \frac{1}{T} \int_0^\tau \mathrm{d}t \; F(x,t)\, \dot{x}, \tag{2.58}$$

where F combines the effect from the potential V and external force f_{ext} and is given by $F = -\frac{\partial V}{\partial x} + f_{ext}$. This equation gives a theoretical expression on the local entropy production of the medium which our Markov model has to replicate.

Since it is not possible to compute equation 2.58, we can rewrite it in terms of the discretized time and space of the Markov state model:

$$\frac{1}{T} \sum_{i=1}^{N} \left(-\frac{V_i - V_{i-1}}{x_i - x_{i-1}} + f_{ext} \right) \frac{x_i - x_{i-1}}{\Delta t} \Delta t = \frac{1}{T} \sum_{i=1}^{N} \left(V_{i-1} - V_i \right) + f_{ext} \left(x_i - x_{i-1} \right), \tag{2.59}$$

where N is a factor relating the lag time τ and the time step of the simulation Δt: $\tau = N\Delta t$. In our model f_{ext} is the same in all microstates. We will make use of the assumption that paths from i to j always follow the same overall direction (possibly against the external force). For big enough systems this assumption holds true, even for periodic boundary conditions. Therefore, all paths in between i and j give us the same result in equation 2.58 and we can use the following relation for the transition probabilities of our Markov model. Performing the sum in equation 2.59 yields:

$$\log \frac{p_{ij}(\tau)}{p_{ji}(\tau)} = \Delta s_{ij} = \frac{1}{T} \left(V_0 - V_N + f_{ext} \left(x_N - x_0 \right) \right), \tag{2.60}$$

since all other terms (with indices not equal to 0 or N) cancel each other. The indices 0 and N describes the potential and the position of state i and j (if we consider the transition $i \to j$) respectively. This discretized version gives us theoretical description of the local entropy production of the medium Δs_{ij} independent of the transition matrix, and therefore, the Markov state model. The theoretical value on the right hand side can now be compared to the left hand side value of the Markov state model.

Chapter 3

Implementation and Model Building

While the Simulation is written in C++, the entire of the following part was implemented in Python.

3.1 Performing the Simulation

The simulations which yield the count matrices to build a Markov state model were performed like described in the theory part in section 2.1 for a single particle in a two dimensional potential.[1] It is essential to get an understanding on the parameters involved and their effect on the simulation.

The equation of motion given in 2.5 is dependent on the external potential V, the external force f_{ext}, the thermal energy k_BT as well as an integration time step Δt. In the following we will disregard all units except for the time. To allow an easy comparison between different simulations, we set the last three parameters constant for all simulations: $k_BT \equiv 5$ and $\Delta t \equiv 10^{-5}$s.

Our potential V consists of three Gaussian wells in a two dimensional box $[0,3] \times [0,1]$ with periodic boundary conditions. These Gaussian wells are symmetric in both directions $x \in [0,3]$ and $y \in [0,1]$ with standard deviations of $\sigma = 0.2$. They are centered at the positions $(0.5, 0.5)$, $(1.5, 0.5)$ and $(2.5, 0.5)$ so they are equally distributed in the configuration space. The depths of these wells, indexed by i from 1 to 3 in ascending order, are of the order $\Delta V_i \sim k_BT$ but can be varied and are not fixed. This way, we can analyze symmetric potentials where each well has the same depth, and also asymmetric ones where the potential depth can vary for different wells (see figure 3.1 for two examples). For the symmetric potentials we choose depths ranging from 1 to 5 $\Delta V_i \in \{1,2,3,4,5\}k_BT$. For the asymmetric ones we set $\Delta V_1 = 1$, $\Delta V_2 = 1$ and vary ΔV_3, again from 1 to 5. In the analysis part we will mainly focus on the two potentials shown in figure 3.1: The smallest symmetric one and the asymmetric one with a $\Delta V_3 = 1$.

[1] It has to be noted that I did not perform the simulations myself. They were provided by Marius Bause, a PhD student in the group.

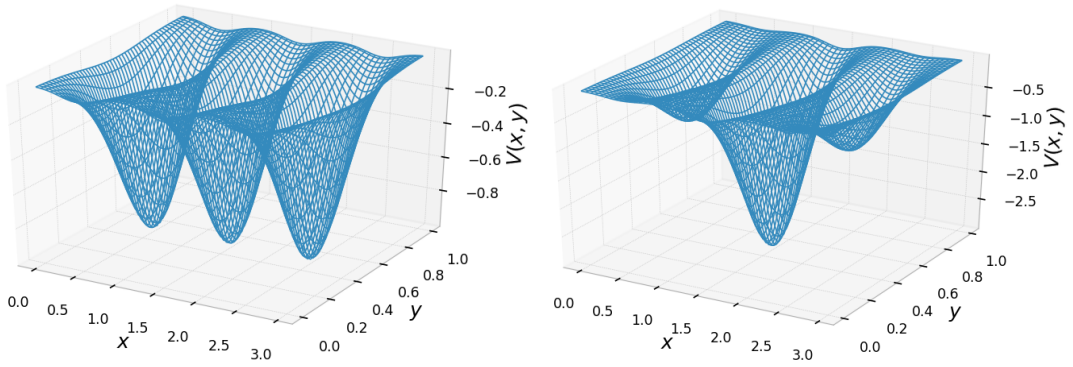

FIGURE 3.1: Left: A Symmetric potential. Right: An asymmetric potential. For illustrative purposes he x- and y-axis scales do not match with each other, in reality the individual wells are symmetric in both directions.

If we apply an external force f_{ext} in the x direction and observe a transition from one well into the other, we want to be able to tell which way the particle took, respectively, if it moved with or against the direction of the force. The smallest system in which this is possible, is a system with three potential wells. For bigger forces this is actually not true anymore and the probability that the particle goes over two wells in the direction of the force is bigger than only jumping one well against it. However, if we compare different strengths of the external force, we can tell which of these processes dominates. The force itself has varying strengths in integer steps from 0 (equilibrium case) to 9 (far of equilibrium) and is pointed in negative x direction. To set the strengths of these forces into context: The greatest force the particle can experience through the external potential is $\sim 6\Delta V_i$.

Since the trajectories consume a substantial amount of storage space, we want to build the count matrices while performing the simulation. In order to do this, we need a space discretization. In our simple system a partitioning of the configuration space into 30 x 10 square boxes of equal size was shown to be sufficient to capture the dynamics. Building the count matrix is simply a matter of counting the transitions, sampled at the lag time $\tau = n\Delta t,\ n \in \mathbb{N}$, between these states. Since we do not know a suitable lag time for a Markov state model at this point, we capture count matrices for lag times between $1\Delta t$ and $999\Delta t$.

3.2 Building the Markov State Model

The estimation of the transition matrix $\mathbf{T}(\tau)$ is straightforward using equation 2.30. We now need to choose a suitable lag time to build the model which we want to use later on. Since the timescale analysis described in section 2.2.2.2 is only possible for reversible transition matrices, we have to look at an equilibrium system. We will assume that the lag time we obtain from this system can also be used for the non equilibrium steady state models. But since we can not have complete certainty about it, we take a conservative estimate of the lag time to leave some room for error. The

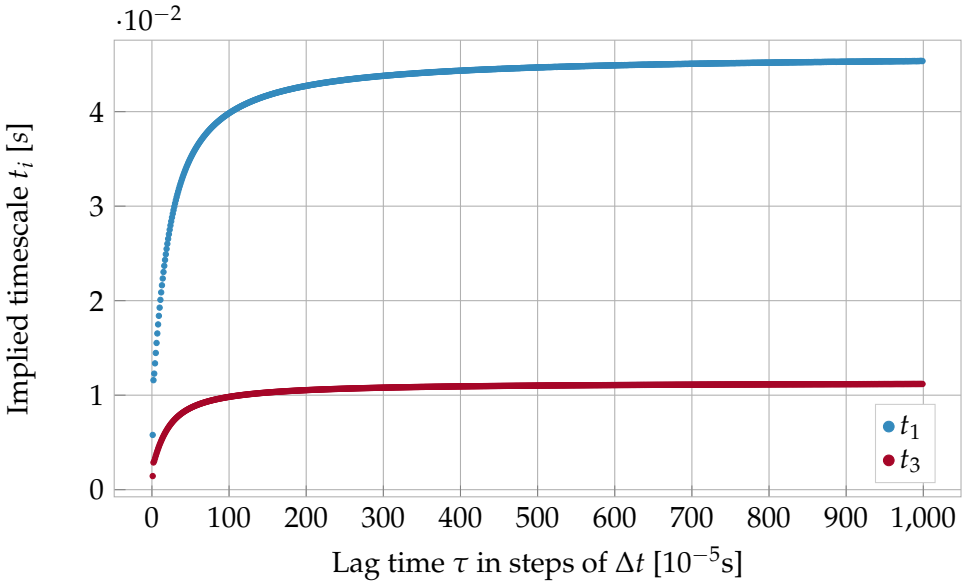

FIGURE 3.2: Convergence of the slowest implied timescales (t_2 is the same as t_1), calculated for the symmetric potential with depths equal to one. For lag times $\tau > 300$ we do not see any significant changes in the timescales anymore so this is the lag time we will choose to build our model.

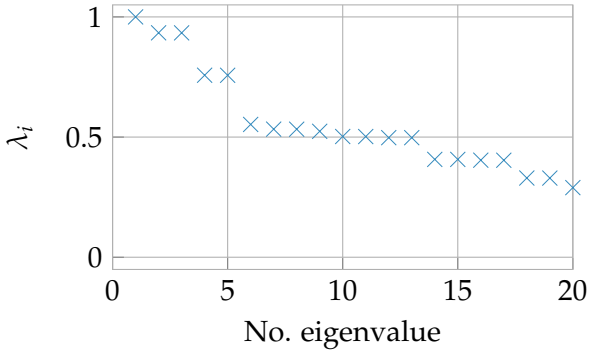

FIGURE 3.3: The first 20 eigenvalues for the symmetric potentials with depths equal to one. The eigenvalues not equal to one always occur in pairs.

slowest implied timescales can be calculated by the highest eigenvalues not equal to 1 with equation 2.20. t_0, with the eigenvalue equal to one, belongs to the stationary distribution and therefore the timescale becomes infinite. Now we need to compare these timescales for different lag times (figure 3.2). The lag time we will use for the model can be determined by looking for convergence in the timescale-lag time plot.

These timescale plots are remarkably similar for all different potentials we use, both symmetric and asymmetric ones. From now on we will use the same lag time of $\tau = 300\Delta t = 3$ms for all these systems. This also allows an easy comparison.

It is worth mentioning that the two biggest eigenvalues, and therefore the two slowest timescales, are almost the same, even for asymmetric potentials. This can be explained by the choice of our potential. In order to be able to represent all transition between the three wells, a superposition composed of two processes is necessary. This effect also continues for faster processes as can be observed in figure 3.3. The eigenvalues always appear as pairs. In the case of the asymmetric potentials, this effect is not as pronounced but nonetheless clearly noticeable.

In order to validate our model, we will now conduct the Chapman-Kolmogorov test (equation 2.34). We will do this by following the implementation suggested in J.H. Prinz et al. [4]. In this test, we want to compare the probability of being in a given set of states A, when starting from a well defined starting distribution. As this starting set A we will choose the ten closest states to a potential minimum in order to represent a metastable set. The corresponding stationary distribution restricted to this set A is then given by:

$$w_i^A = \begin{cases} \frac{\pi_i}{\sum_{j\in A}\pi_j} & \text{for } i \in A \\ 0 & \text{otherwise,} \end{cases} \tag{3.1}$$

where π_i is the probability to be in state i.

We will now compute the probability of finding the system in set A at times $k\tau,\ k \in \mathbb{N}$: $p\,(A, A; k\tau)$. Starting with the initial probability vector $\mathbf{w}^A$, we will use the observed trajectory data, as well as our Markov state model in order to do this. If the probabilities calculated by the trajectory data and the Markov model conform with each other, we essentially confirmed the Chapman-Kolmogorov equation.

The probability to be in at set A according to the Markov model is given by:

$$p_{Markov}\,(A, A; k\tau) = \sum_{i\in A}[\left(\mathbf{w}^A\right)^T \mathbf{T}^k\,(\tau)]_i\,, \tag{3.2}$$

where $\mathbf{T}\,(\tau)$ is the transition matrix of our Markov state model.

On the other hand is the trajectory based probability given by:

$$p_{Traj}\,(A, A; k\tau) = \sum_{i\in A} w_i^A \frac{\sum_{j\in A} c_{ij}\,(k\tau)}{\sum_{j=1}^{n} c_{ij}\,(k\tau)}, \tag{3.3}$$

where n is the total number of states and c_{ij} the entries of the count matrix obtained from the trajectory for transitions at every step $k\tau$. The fraction in this equation describes the probability to be in set A at time $k\tau$ when starting in state i.

Since there are only a finite number of transitions available to estimate these probabilities, there will be statistical error involved. To account for this, the uncertainties (one sigma standard error) of the probabilities estimated from the trajectories are calculated using the following equation:

$$err_{Traj}\,(A, A; k\tau) = \sqrt{k\frac{p_{Traj}\,(A, A; k\tau) - [p_{Traj}\,(A, A; k\tau)]^2}{\sum_{i\in A}\sum_{j=1}^{n} c_{ij}\,(k\tau)}} \tag{3.4}$$

The test will now consist of whether the two probabilities confirm with each other within this uncertainty. In figure 3.4 this was done for several k-values (for bigger values the statistical uncertainty gets too large). We used the Markov model and a simulation trajectory for the smallest symmetric potential (figure 3.1 on the left) and the lag time we obtained from the timescale analysis of $\tau = 300\Delta t$.

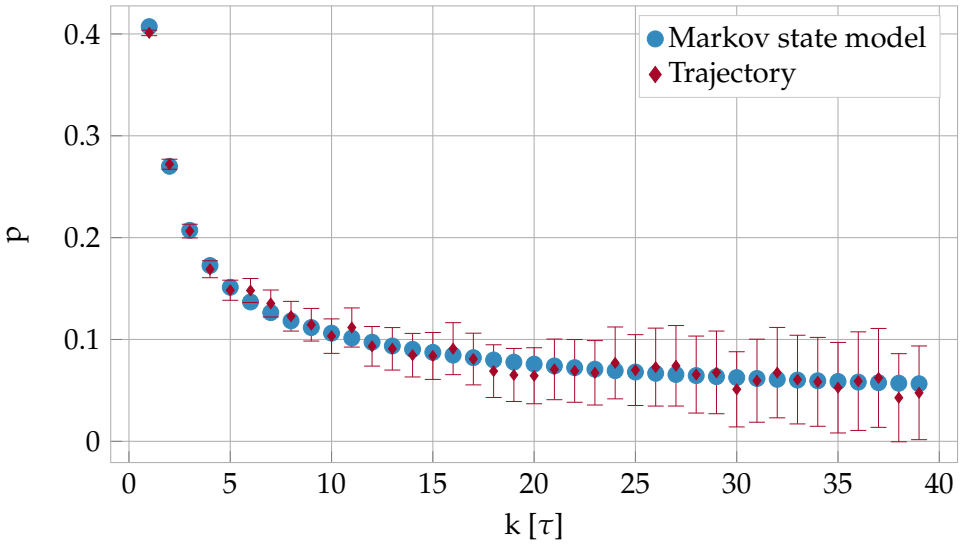

FIGURE 3.4: Chapman-Kolmogorov test performed for the symmetric potential and a lag time of $\tau = 300\Delta t$. The probabilities to stay in a potential minimum A are calculated by the Markov model and directly from the trajectory itself. Within their statistical error these two values conform with each other.

As we can see in figure 3.4, both values conform with each other and we have therefore validated the Markov state model built with lag time $\tau = 300\Delta t$. This test was only performed for this specific case. Since the timescale plots for different potentials did not vary much for different potentials, we will assume that the Chapman-Kolmogorov equation will still hold true in these cases.

We have discretized space and time in order to capture the timescales of interest. Now that we have Markov state models for all our desired systems, we can start with the reweighting process.

3.3 Reweighting

3.3.1 Legitimation of the Average Flux

First we want to legitimize our choice for the average flux $\langle J \rangle$ defined in equation 2.46.

In order to do this, we need to define the matrix F_{ij} which depends on the geometry of the system. In our case the external force only acts in the x direction, so all entries for transitions in y direction are zero. For every step in the x direction we need to add $\pm 1/10$ (the size of one microstate), dependent on the direction of the transition in question. This way our matrix has entries ranging from -1.4 to 1.4 since we can go 14 steps in each direction. If we go 15 steps we do not know which direction the transition took, so we set it to zero. However, the transition matrix shows transitions in a range of 1-9 states with our choice of microstates, potential and external force (figure 3.5). All the other ones are so unlikely that they did not happen once in the sampled trajectory. This way we always know the direction the

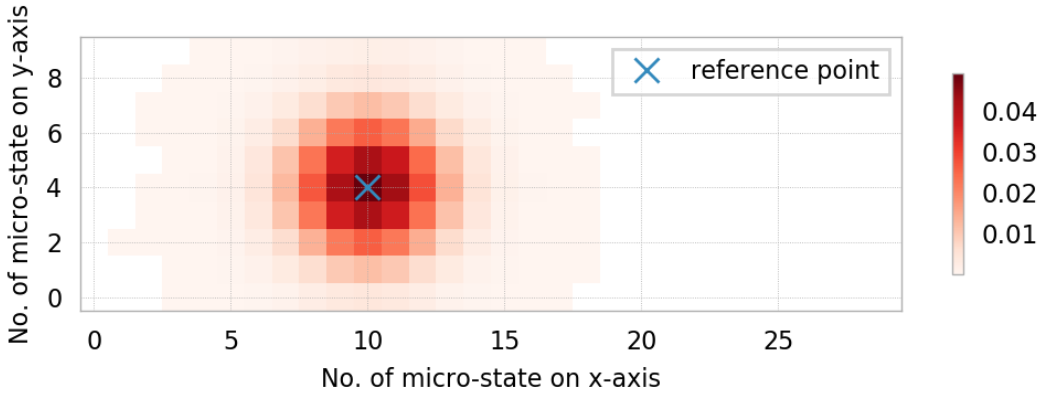

FIGURE 3.5: The transition matrix is sketched for a single reference point i on top of a potential well to allow the farthest transitions.

transition has taken. The average flux for systems with different external forces plotted against the lag time (figure 3.6), shows with great certainty that the average flux describes a Markovian macroscopic observable.[2] This legitimizes our definition for the average flux $\langle J \rangle$ and the matrix F_{ij}.

Furthermore, there seems to be a direct relation between the average flux and the external force. This relation can be obtained directly from the overdamped Langevin equation 2.5 by averaging over time:

$$0 = -\langle \frac{\partial V}{\partial x} \rangle_t + \langle f_{ext} \rangle_t - \langle v \rangle_t + \langle f_R \rangle_t. \tag{3.5}$$

In this equation $\langle \rangle_t$ indicates the averaging over time. From $\langle \frac{\partial V}{\partial x} \rangle_t$ we obtain some number ϵ caused by the perturbation from the external force. The force on the other hand is constant in time, so $\langle f_{ext} \rangle_t = -f_{ext}$ (the $-$ sign comes from the direction

[2]Furthermore, the average flux was calculated directly from a trajectory. This value was in conjunction with the one determined by the Markov model.

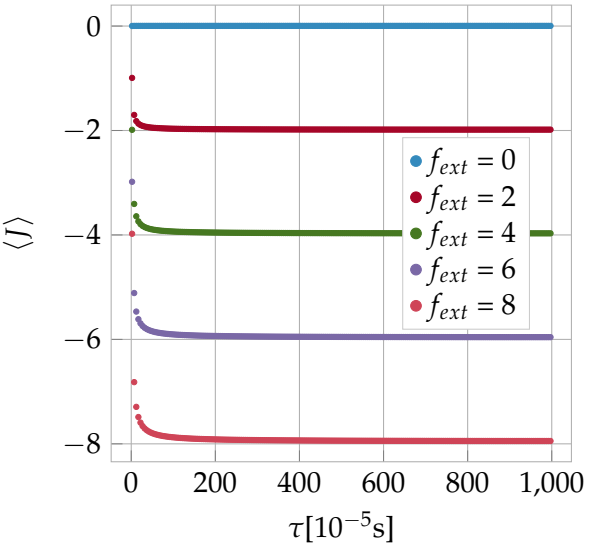

FIGURE 3.6: The average flux $\langle J \rangle$ describes a Markovian behavior for high enough lag times τ. We also see the direct relation between the average flux and the external force fairly well.

of the force). The temporal mean of the velocity corresponds to our average flux $\langle v \rangle_t = \langle J \rangle$ and the mean of the random force is zero by definition. From this we can set up the following equation:

$$\langle J \rangle = -f_{ext} - \epsilon. \tag{3.6}$$

By our observation in figure 3.6 we conclude that epsilon is some small number since the average flux deviates only by a very small amount, which slowly increases with increasing force.

3.3.2 Reweighting Algorithm

We have derived an analytic expression for the Markov state model p_{ij} of a desired system, characterized by its flux $\langle J \rangle$, given a reference model q_{ij} (equation 2.52):

Here we test the reweighting between models which have the same potential but are in different non equilibrium steady states. The reference system does not need to be in equilibrium.

The only challenge remaining is to determine the Lagrange multiplier γ, controlling the $\langle J \rangle$ part in the Caliber. In order to this, we first look at the new calculated average flux with varying γ given the same reference system:

$$\langle J \rangle_{rew} (\gamma) = \sum_{ij} \pi_i (\gamma) \, p_{ij} (\gamma) \, F_{ij}. \tag{3.7}$$

The results are depicted in figure 3.7. Whereby the reference system in this particular case is the system with symmetric potential in the equilibrium state (therefore we have a flux of 0 for γ of zero). We conclude that the reweighted average flux is strictly monotonically increasing in γ. This is an important realization since in order to determine our new model $p_{ij} (\gamma)$, we need to minimize the following function:

$$\left(\langle J \rangle_{rew} (\gamma) - \langle J \rangle_{desired} \right)^2. \tag{3.8}$$

Because of the strictly monotonically behavior of the reweighted flux, this problem has exactly one solution and can be solved by a simple downhill simplex algorithm [27]. The desired flux $\langle J \rangle_{desired}$ will be obtained by a Markov model, built from simulation outcome, in the same steady state.

Another important realization to be made is that the reweighted flux shows a strictly linear behavior in γ for values of interest to us (figure 3.7 on the right). This means that we only have to conduct the reweighting algorithm one or two (to be sure) times in order to get a relation between the average flux and γ. This is very beneficial for us since it allows us to determine the Markov state model for any desired system, in a certain range, instantly without having to minimize equation 3.8. This reduces the time needed to perform a reweighting significantly.

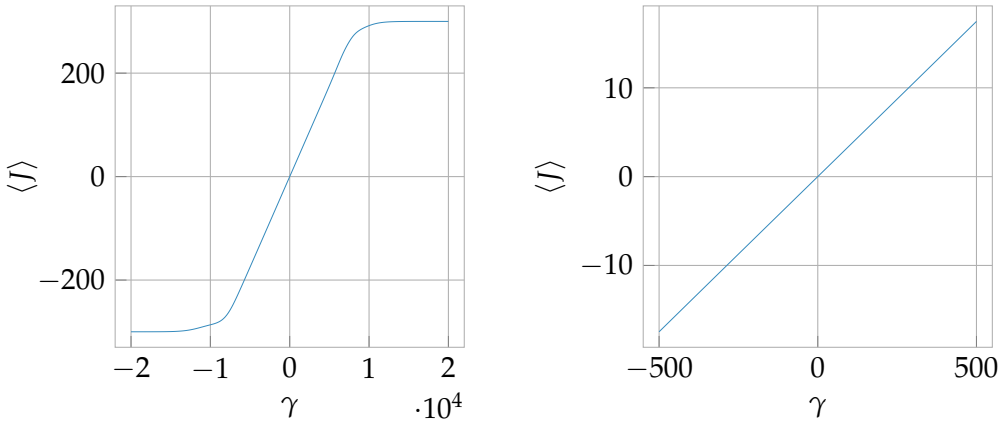

FIGURE 3.7: The reweighted average flux $\langle J \rangle$ plotted against the Lagrange multiplier γ. On the left we see the bigger picture: $\langle J \rangle$ is strictly monotonically increasing in γ (shows an arctan behavior). On the right we have a zoomed in plot of the for us interesting area. Here $\langle J \rangle$ shows a linear behavior which we can use to our advantage.

3.4 Calculating the First Passage Time Distribution

In order to access the quality of the reweighting process we take a look at the first passage time distributions. These can be obtained as described in section 2.4.1.

In our system the most interesting first passage times are the ones between the different potential wells. Therefore, we need to assign microstates to these minima.

These can be obtained by determining each local minimum in the non discretized space. By defining a threshold κ we can assign an area in this space to each minimum using the following condition: $V(x,y) < -\Delta V_i + \kappa$,where ΔV_i describes the potential depth. At this point we can search for the microstates which overlap with this area and assign them to these minima. In our case we get four microstates for each minimum (with a κ value of $\Delta V_i/10$) where each minimum is located at an intersection between these four states.

So in order to calculated the first passage time distributions, we need to expand our formula (equation 2.55) for different sets A and B: First we redefine $f_{ij}^{(n)}$:

$$f_{ij}^{(n)} = \begin{cases} p_{ij} & for \;\; n = 1 \\ \sum_{m \notin B} p_{im} f_{mj}^{(n-1)} & for \;\; n > 1 \end{cases} \tag{3.9}$$

This way we do not take pathways into account which already visited the ending set B. Now we need to sum over all states in the starting set A and the ending set B and weight the starting states with the probabilities that the system will start in these states. These weights are defined as follows:

$$w_i = \sum_{k \notin A} \pi_k p_{ki} \tag{3.10}$$

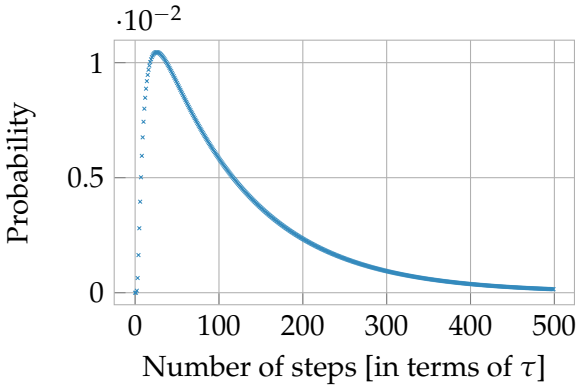

FIGURE 3.8: An example for a first passage time distribution, from one local minima into another in the smallest symmetric potential in equilibrium.

Note that we do not use the stationary distribution as weights. We do need to account for the fact that our system is moving. Therefore, we are calculating the probability of the starting state being entered. Our equation for the first passage time distribution becomes,

$$f_{AB}^{(n)} = \sum_{i \in A, j \in B} \frac{w_i}{\sum_{i \in A} w_i} f_{ij}^{(n)}. \tag{3.11}$$

Calculating this for several n gives us a probability distribution of the first passage times for transitions from set A to set B in steps of the lag time τ. An example for such a distribution is given in figure 3.8. From this distribution we can calculate the first three moments: The mean value, the standard deviation and the skewness. These moments allow for an easier comparison of different distributions.

Chapter 4

Analysis

4.1 Mean First Passage Time

First we want to have a look at the first passage time distribution for different lag times τ. We can calculate these by using the formula we derived in section 3.4. In order to compare the distribution we will have a look at the mean first passage times. If we plot the mean first passage times for different lag times (figure 4.1 on the left) we observe a non Markovian behavior. At first (until $\tau \sim 200$) the discretization error of the Markov model dominates. This is the same behavior we also observe during the timescale analysis for different lag times. However, even after this initial part there is a steady increase for increasing lag times. We conclude that the mean first passage time is not a Markovian observable.

This is rather surprising since the mean first passage time is often used to describe the dynamical properties of Markov state models [11]. Nevertheless, the first passage time distribution gives the best description of a process for a chosen lag time while the timescale analysis is not available off equilibrium [28].

In the following we will stick to the lag time of $\tau = 300$ in order to be able to compare different first passage time distributions with each other.

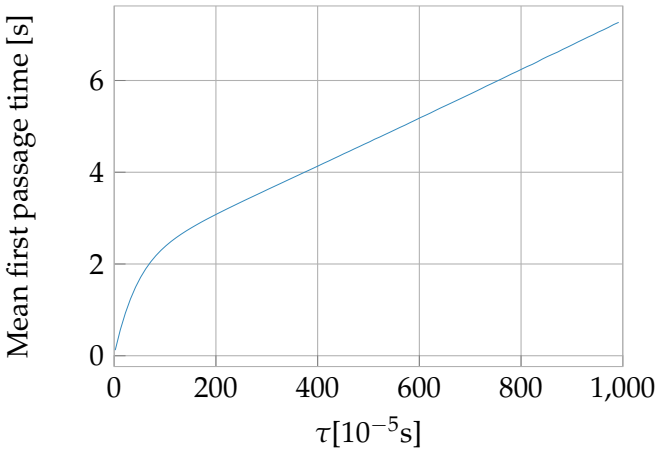

FIGURE 4.1: The mean first passage time does not show a Markovian behaviour.

4.2 Shift of the Stationary Distribution

If we take a look at the stationary distribution, one would expect that it would shift inside the minima according to the direction of the force (particle gets pushed to the side in the potential well). However, the exact opposite effect occurs. In figure 4.2 we see the stationary distribution in the full discretized state space. The author can not give an explanation on this behavior at this point and further research needs to be done.

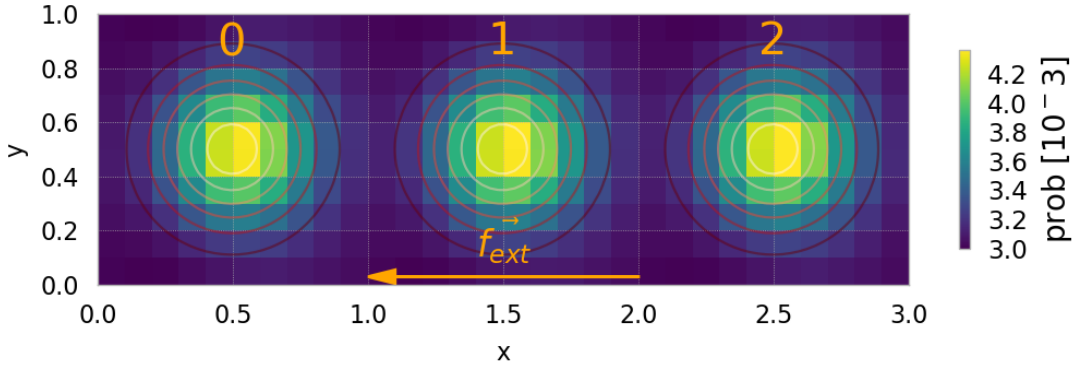

FIGURE 4.2: Stationary distribution for the smallest symmetric potential. The arrow indicates the direction of the external force and one can clearly see that the occupation probabilities shift in the opposite direction. Furthermore we illustrate the indexing for each minimum.

4.3 Reweighting Tested for Symmetric Potentials

In this section we will first have a look at the reweighted Markov state models for the smallest symmetric potential (the one where all potential wells have the same depth $\Delta V_i = 1$, figure 3.1 on the left).

4.3.1 Predicting First Passage Time Distributions

In figure 4.3 we see mean first passage times for transitions with or against the direction of the external force, obtained from models directly build out of the simulation outcome. In order to determine error bars this was done for 20 different count matrices.

The behavior of the two transitions is different depending on the direction the transition is taking. The plot on the left of figure 4.3 shows the mean first passage times for a transition in the direction of the external force. As we would expect in this case, the mean first passage time is linearly decreasing with an increasing force f_{ext}. The particle gets pushed in the direction of the transition we are investigating and therefore the mean first passage time decreases.

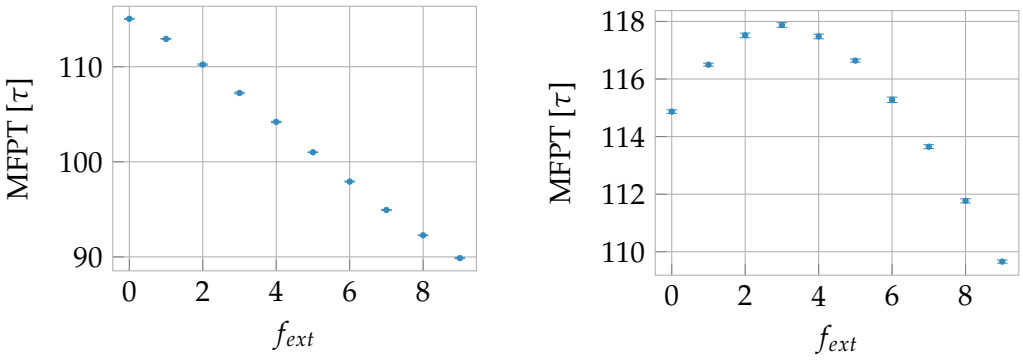

FIGURE 4.3: Mean first passage times (MFPT) for different strengths of the external force. On the left a transition in the direction of the external force is depicted, on the right the transition against the direction of the force.

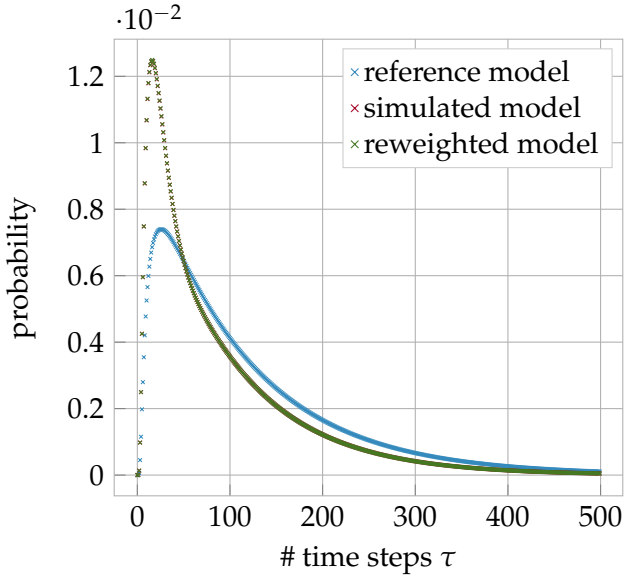

FIGURE 4.4: Three first passage time distributions for a transition in the direction of the force. The reweighted model is obtained by the reference model and matches the outcome of the desired model perfectly. The reweighted data points are right on top of the simulated ones.

In the right plot we are looking at the transition in opposite direction. First the mean first passage time is increasing because the particle has to go against the direction of the force. For larger forces however, it is decreasing again. The explanation for this is straightforward: For larger forces we are actually observing the transition in the direction of the force over two wells. Different processes dominate for different external forces.

Now we will take a look at the first passage time distribution for a reweighted Markov state model 3.3.2. We will reweight the equilibrium system into the system out of equilibrium with highest external force $f_{ext} = 9$. In figure 4.4 the distributions are depicted for the models in and out of equilibrium obtained by a simulation, as well as the first passage time distribution obtained by the reweighted model. In this case the error bars are not included since they were smaller than the point size.

As we can see, in this case the reweighting accurately reproduced the expected behavior. But we might want to know what happens if we take a system out of equilibrium as our reference model.

Turning to these cases in figure 4.5 we take an off equilibrium reference model

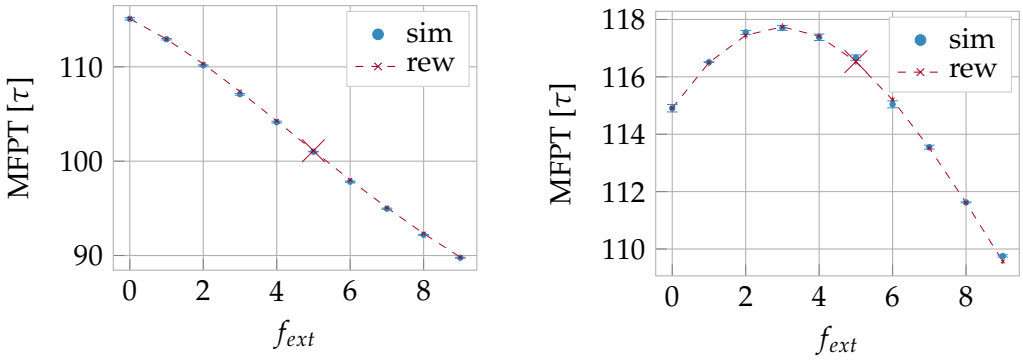

FIGURE 4.5: Mean first passage times of the reweighted models (red) compared to the ones calculated from the simulation models (blue). On the left we go in the direction of the external force, on the right against it. The dashed line illustrates that we can reach all these areas with a suitable γ. We reweight from the system with $f_{ext} = 5$ in these plots which is indicated by a bigger cross.

with an external force $f_{ext} = 5$ that is reweighted into every other system in its vicinity. We compare the mean first passage times, both from the simulated and reweighted models, for transitions in both directions.

All the reweighted models predicted the mean first passage times of the simulated models. These results are independent of the reference model we use. Furthermore, when looking at higher moments we get the same results (data is not shown in this thesis).

Apparently the reweighting successfully predicts first passage time distributions between different minima. Now we will take a look at the entropy production.

4.3.2 Predicting the Entropy Production Rate

In order to estimate the entropy production rate, as well as the local entropy production, from our Markov state models we use the equations derived in section 2.4.2.

First we look at the entropy production rate for simulated and reweighted models. As a reference model we take the equilibrium system, reweight it into every available steady state and compare it with the simulated model. As one can see in the upper plot in figure 4.6 the entropy production for the reweighted and simulated models fit perfectly. Again errors are not taken into account since they are not noticeable in the plot.

Secondly we want to discuss the local entropy production (figure 4.6 bottom): We take the difference between the theoretical value (equation 2.60) and the one calculated by our model (equation 2.57). For the starting microstate we chose a state in the middle of a potential well. The simulated model as well as the reweighted model are set in the system with the highest external force $f_{ext} = 9$ while the equilibrium system was used as a reference model. These differences are in the range of around 10% (deep red/blue) of the theoretical value. The transitions with larger distance from the starting microstate suffer from poor sampling and, therefore, have a higher

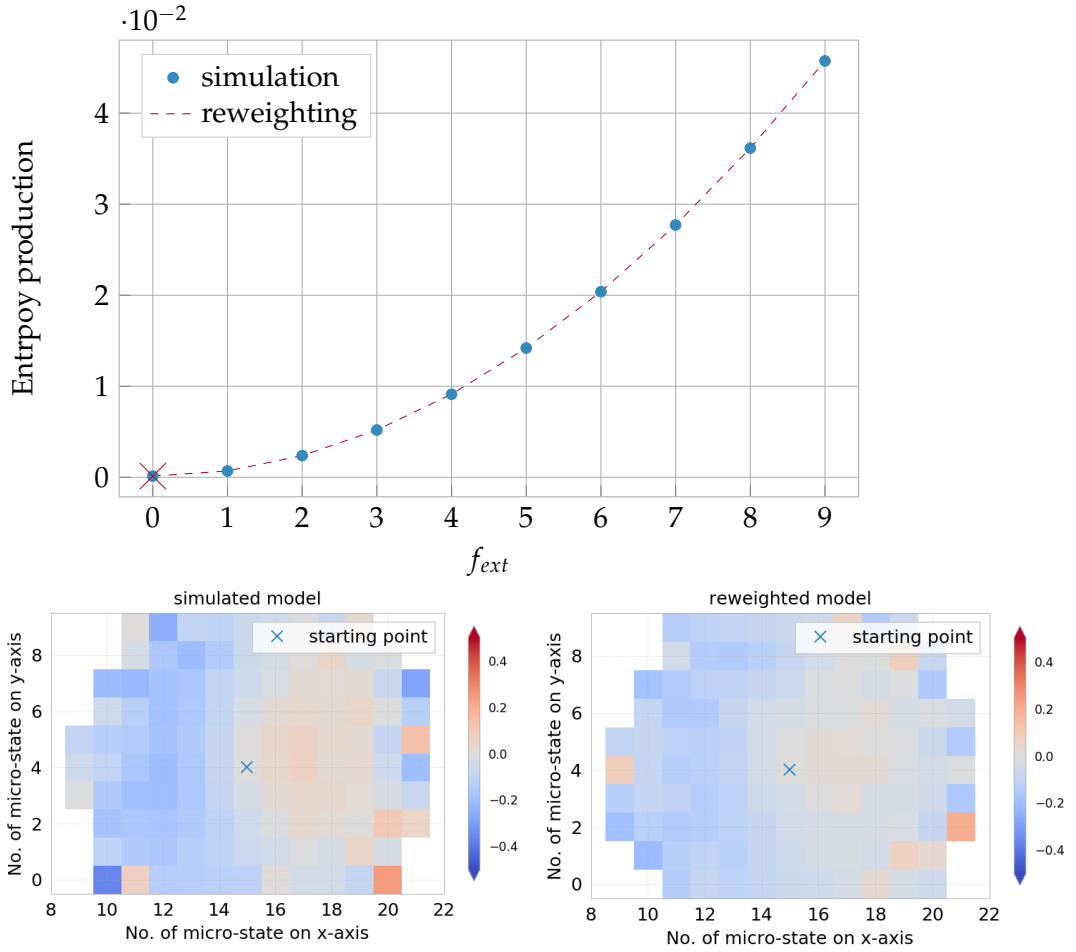

FIGURE 4.6: Top figure: The Entropy production rate for each simulated model and the reweighted one (reweighted from equilibrium). They fit perfectly. Bottom two figures: Illustration of the difference between the theoretical local entropy production and the one calculated from the Markov model. Both for the model obtained from a simulation (top) and from reweighting (lower). There are no major differences between the two models.

deviance and are not representable.[1] In general (also taken other systems and reference systems into account which are not illustrated here) it can be said both models recover the theoretical description of the local entropy production. We can draw two conclusions from this: First the assumption we make in the theory part 2.4.2.1, replacing a single trajectory by all possible paths described by our Markov model, can be validated. Furthermore, this local entropy production allows us to have a detailed look at the dynamics of the system which get recovered by the reweighting process.

Again, these results are independent of the reference model we used.

4.3.3 Stationary Distributions

The stationary distribution of our two dimensional model is defined as the probability to find our system at a certain microstate for each state in the x-y plane. In order

[1] In order to calculate these we build in a cutoff transition probability so the logarithm does not diverge. These are already taken out of the plot in figure 4.6

to allow for an easy comparison of stationary distributions for different non equilibrium steady state systems, we project the distribution onto the x-probability plane. In this case we can observe how the stationary distribution is changing along the x-axis. We do not expect any change along the y-axis for different driven systems.

In figure 4.7 we compare the stationary distributions of the reference model, the simulated model we want to reweight to and the reweighted model. We do this once for a reweighting from an equilibrium system in a far off equilibrium system ($f_{ext} = 9$) (left plot) and once the other way around (right plot).

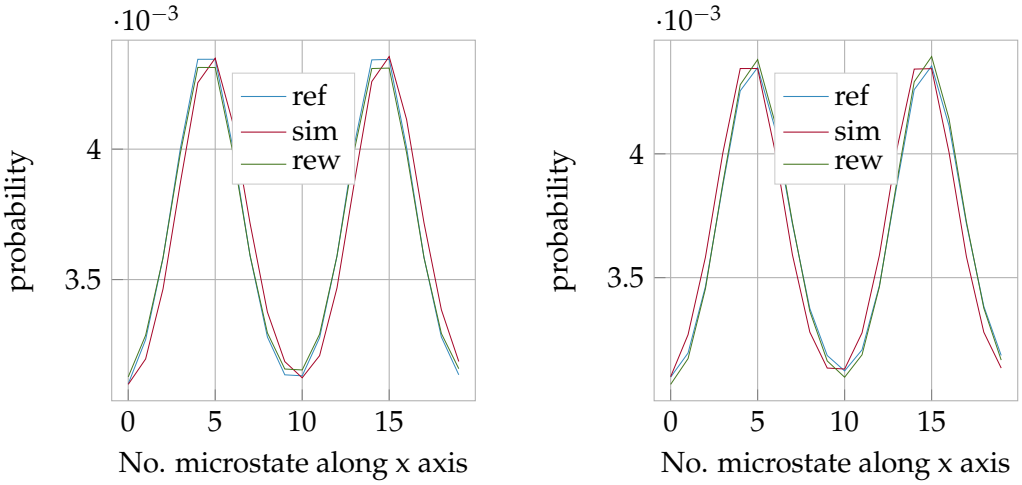

FIGURE 4.7: Comparison of the stationary distribution of a reference model (blue), a simulated model (the one we will reweight in, red) and a reweighted model (green). On the left: Reweighting from an equilibrium system into a non equilibrium one. On the right: Reweighting from a non equilibrium system into an equilibrium one. The reweighted models are not able to reproduce the stationary distributions of the systems in question.

We observe that the stationary distribution of the reweighted model is very similar to the one from the reference model. It does not adapt to the asymmetry of the non-equilibrium system (or keeps it, if the reference model is the non-equilibrium one).

However, these small deviations in the stationary distributions do not seem to have an influence on the dynamics of the system as was shown by the first passage times and the entropy production rate.

The results we obtained in this section are independent on the potential well depth (data not shown in this thesis) as long as we deal with a symmetric potential.

4.4 Reweighting Tested for Asymmetric Potentials

We access the quality of the reweighting method the same way we did for the symmetric potentials but starting with the stationary distribution. In order to do this we have a look at an asymmetric potential where two wells have the same depths ($\Delta V_i = 1$) and one (the 'middle' one) is deeper than the other two ($\Delta V_1 = 3$) (see figure 3.1 on the right).

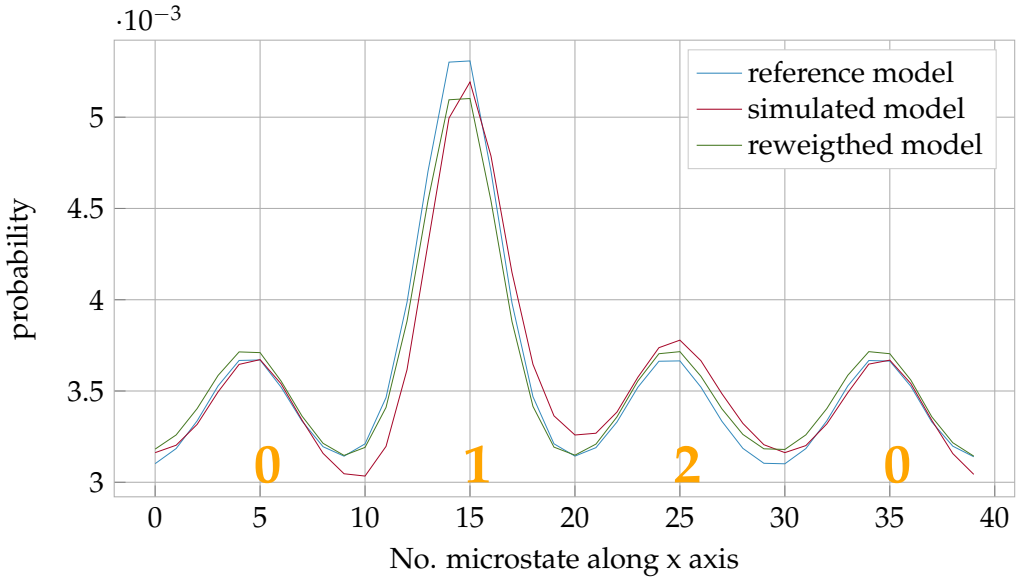

FIGURE 4.8: Comparison of the stationary distribution of a reference model (blue), a simulated model (the one we will reweight in, red) and a reweighted model (green). The numbers indicate the indices of the potential wells. After microstate 30 the system crosses the periodic boundary conditions and start in well 0 again.

In figure 4.8 the stationary distribution is shown for a reweighting from an equilibrium system to a system far off equilibrium (again $f_{ext} = 9$). It is compared to the stationary distribution of the reference model as well as the one of the non equilibrium simulated model.

The same effect as for the symmetric potential occurs: The stationary distribution shape is similar to the one of the reference system. Not only does it disregard the shift of the probability caused by the external force in each peak but also changes in the amplitude. Especially the probability of finding our system in between two potential wells is not recovered correctly. In between the wells 0-1 and 1-2 we find significant deviations.

4.4.1 Dynamics

We will look at the same dynamical properties as for the symmetric potential and see if the reweighted model still predicts these correctly.

The entropy production rate (figure 4.9 top) is correctly given by the reweighted model. Furthermore, we see that also the local entropy production is correctly recovered, even for the area where the stationary distributions show a significant deviation (figure 4.9 bottom).

However, the reweighted model can no longer predict the mean first passage times between different potential wells anymore (see figure 4.10). We assume systematic error to be involved, causing the reweighted model to either over- or underestimate a certain transition. In figure 4.10 we look at the transitions $0 \to 1$ (on the left) and $1 \to 2$ (on the right) where 1 is the deep potential well. The equilibrium system is chosen as the reference system. Comparing these errors with the errors in

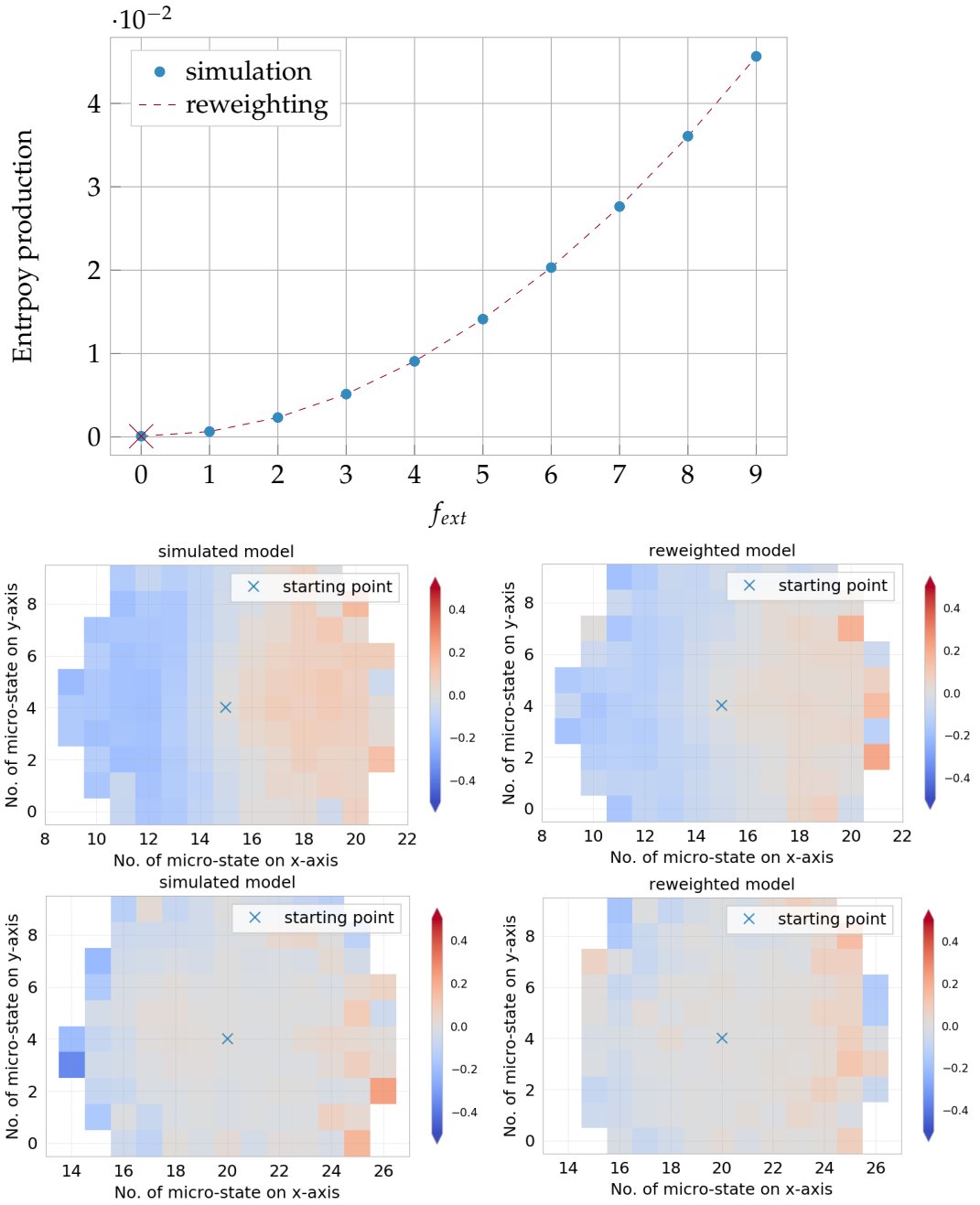

FIGURE 4.9: Top figure: The Entropy production rate for each simulated model and the reweighted one (reweighted from equilibrium). They fit perfectly. Bottom four figures: Illustration of the difference between the theoretical local entropy production and the one calculated from the Markov model for different starting points. Both for the model obtained from a simulation (left) and from reweighting (right). Here the upper two plots are for the starting point within a potential well (same as in fig 4.6) and the lower two for a starting point in between two wells where the stationary distribution is not recovered correctly.

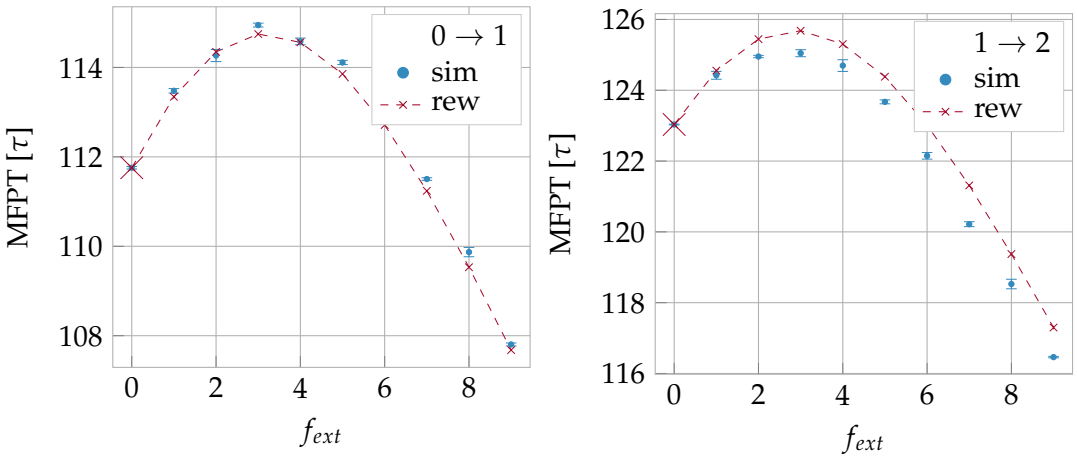

FIGURE 4.10: Mean first passage times of the reweighted models (red) compared to the ones calculated from the simulation models (blue). On the left we see the mean first passage times for the $0 \to 1$ transition, in the right one the $1 \to 2$ transition. The equilibrium model is chosen to be the reference model in these plots indicated by the bigger cross.

the stationary distributions leads us to the conclusion that they are correlated. For the transition $0 \to 1$ the reweighted mean first passage times are too fast. At the same time the probabilities of our system to be in between these two wells (figure 4.8) is overestimated. The reweighted system is more likely to be in states between 0 and 1 and therefore the transition occurring between these two is a faster process. In the second case for the $1 \to 2$ transition it is the other way around. The reweighted mean first passage time is too slow and the probability in between underestimated.

All the other transitions are illustrated in figure 4.11. For the $1 \to 0$ transition the same explanation as for the $0 \to 1$ one can be given. However, for the transitions starting in the potential well 2 the explanation is not as obvious anymore since the stationary distribution also shows a deviation inside this well which affects the transition times as well.

In the case of the symmetric potentials this was no problem since the overall probability of the system to be in a certain potential well, respectively be in between two potential wells, did not change because the error is averaged out when defining sets of microstates in the minima.

The reweighting process with the maximum Caliber approach does not account for the changes in the stationary distribution, at least not enough. For simple systems like the symmetric potential well this does not have any effects on the macroscopic dynamics of the systems. However, for more complex systems the deviations in the stationary distribution do have a significant effect on the dynamics and lead to wrong predictions. The entropy production rate of the whole system is not affected by these deviations. Furthermore, the local entropy production can be recovered as well.

We can conclude the following statements:

- The dynamics of the system as a whole, described by the entropy production

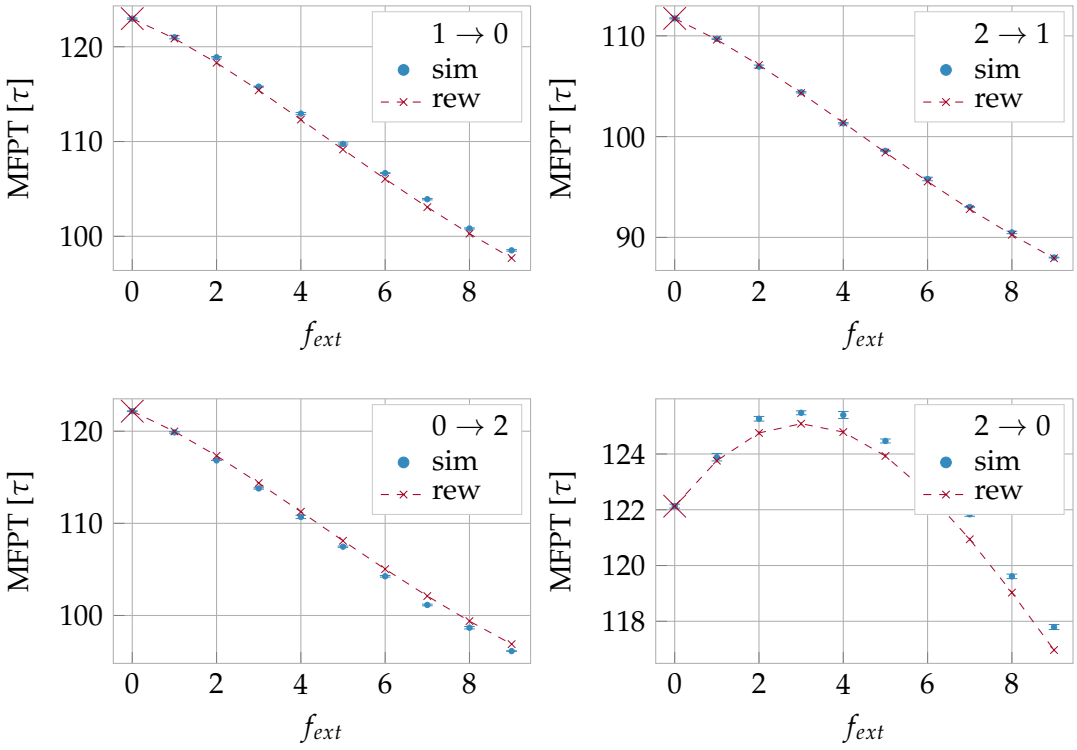

FIGURE 4.11: Mean first passage times for all other transition paths. The reweighted models (red) compared to the ones calculated from the simulation models (blue). The equilibrium model is chosen to be the reference model in these plots indicated by the bigger cross.

rate, can be recovered and are unaffected by the deviations in the stationary distribution.

- Dynamics between potential wells, between different potential wells can not be predicted properly since they get influenced by the false stationary distribution.
- There seems to be some microscopic property that preserves the local entropy production rate, namely the ratio p_{ij}/p_{ji} (Markov model description of the local entropy production: $\Delta s_{ij} = \log\left(p_{ij}/p_{ji}\right)$) is reproduced by the reweighting process.

In order to solve the issue of falsely predicted local dynamics within the system, we need the reweighted Markov state model to correctly predict the stationary distribution. This has to be done by tweaking the Caliber. One way to solve this would be to find a theoretical approach to predict the changes in the stationary distribution. In order to do this, we first need to understand what causes the asymmetry in the stationary distributions on a fundamental level. Another way would be to find another macroscopic property, such as we already have with the average flux, that influences the stationary distribution and can be put into the caliber as another constraint.

Chapter 5

Conclusion and Outlook

In this thesis we tested a reweighting method based on the principle of maximum Caliber applied to Markov state models. This allows to predict Markov state models which normally have to be calculated by a simulation.

An important realization to be made is that the Lagrangian coefficient γ shows a strictly linear behavior to the average flux that we use to describe how far a system is off-equilibrium. This means that we only have to simulate two steady state systems in order to be able to reweight into a broad range of other steady state systems. How far off-equilibrium one ca go with this method is only limited by the periodic boundary conditions: Jumps in the Markov model have to be smaller than half system size in order to still be able to determine their direction.

In order to assess the quality of the reweighted model, we compared dynamical properties like the entropy production rate and the first passage time distributions for reweighted models in different non equilibrium steady states. We discovered that the entropy production rate, describing the dynamics of the whole system can be correctly predicted by the reweighted Markov state models. However, more local dynamics like the first passage time distributions between different potential minima are not necessarily recovered by the reweighting process, dependent on the complexity of the system in question. This is due to the fact that the reweighted models falsely predict a stationary distribution which is too similar to the one of the reference model used for reweighting. This result was verified in a one dimensional potential by implementing the correct stationary distribution as a further constraint in the Caliber: The system obtained by this tweaked Caliber recovered the first passage time distributions perfectly, both in the case of a symmetric as well as asymmetric potential. Therefore, we need to find the reasons behind the behavior of the stationary distribution and expand our maximum Caliber accordingly, so it is able to predict the changes in the stationary distribution.

Nonetheless, we can conclude that the reweighting is already working , although only to a certain extend. This allows to save simulation time. The reweighting method also offers the possibility to built models for different non equilibrium steady states in a continuous manner, whereby simulations are only able to model a single

one. For all of that we only need one single characteristic quantity so far, the average flux. These results also support the theory of maximum Caliber and the assumption of the Kullback-Leibler divergence, that reference and the reweighted model are closely related, since the reweighting results are independent of the reference model we use.

We confirmed the relation for the local entropy production between a theoretical value and the one calculated by the Markov state model. This property is conserved in the reweighting process. The assumption to replace a single trajectory by the Markov model is therefore verified for systems with a constant external force. This yields a new relation, similar to the detailed balance one, for non equilibrium steady states:

$$\text{detailed balance: } \frac{p_{ij}}{p_{ji}} = \frac{\pi_j}{\pi_i}, \qquad \text{local entropy production: } \frac{p_{ij}}{p_{ji}} = \exp\left(\Delta s_{ij}\right). \tag{5.1}$$

The same way detailed balance is used to construct Markov state models in equilibrium systems, this relation can be used for the construction of models in non equilibrium steady state systems. This allows to improve the model for transitions which suffer from poor sampling in only one direction. Testing this new relation in the construction of Markov state models and as a constraint in the Caliber is an objective for the future.

Furthermore, we should reconsider the role of the mean first passage time as an observable in a Markov state model. It is not Markovian, but for same lag times we still get comparable results.

In the future the reweighting can be tested for more complex systems, for example multiple particle and polymer systems.

We could also test the reweighting method on coarse grained Markov state models obtained by the clustering method G-PCCA (generalized Perron-cluster cluster analysis) [29] which works for Markov state models breaking detailed balance. New challenges will occur, because the clustering is not necessarily equal for all off-equilibrium systems.

Bibliography

[1] Nuria Plattner and Frank Noé. "Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models". In: *Nature Communications* 6.1 (2015). DOI: 10.1038/ncomms8653.

[2] Stephen J. Klippenstein, Vijay S. Pande, and Donald G. Truhlar. "Chemical Kinetics and Mechanisms of Complex Systems: A Perspective on Recent Theoretical Advances". In: *ChemInform* 45.18 (2014). DOI: 10.1002/chin.201418296.

[3] H. Wan, G. Zhou, and V.A. Voelz. "A Maximum-Caliber Approach to Predicting Perturbed Folding Kinetics Due to Mutations". In: *Journal of Chemical Theory and Computation* 12.12 (2016). DOI: 10.1021/acs.jctc.6b00938.

[4] Jan-Hendrik Prinz et al. "Markov models of molecular kinetics: Generation and validation". In: *The Journal of Chemical Physics* 134.174105 (2011).

[5] Han Wang and Christof Schütte. "Building Markov State Models for Periodically Driven Non-Equilibrium Systems". In: *Journal of Chemical Theory and Computation* 11.4 (2015), pp. 1819–1831. DOI: 10.1021/ct500997y.

[6] Daniel M.Zuckerman. "Non-equilibrium: Steady States". Physical lens on the cell. 2015.

[7] Steve Pressé et al. "Principles of maximum entropy and maximum caliber in statistical physics". In: *Reviews of Modern Physics* 85.3 (2013), 1115–1141. DOI: 10.1103/revmodphys.85.1115.

[8] Purushottam D. Dixit et al. "Perspective: Maximum caliber is a general variational principle for dynamical systems". In: *The Journal of Chemical Physics* 148.1 (2018). DOI: 10.1063/1.5012990.

[9] Udo Seifert. "Stochastic thermodynamics, fluctuation theorems and molecular machines". In: *Reports on Progress in Physics* 75.12 (2012).

[10] J. Schnakenberg. "Network theory of microscopic and macroscopic behavior of master equation systems". In: *Reviews of Modern Physics* 48.4 (1976), 571–585. DOI: 10.1103/revmodphys.48.571.

[11] Ernesto Suárez et al. "Estimating first-passage time distributions from weighted ensemble simulations and non-Markovian analyses". In: *Protein Science* 25.1 (2015), 67–78. DOI: 10.1002/pro.2738.

[12] M. Laleman E. Carlon and S. Nomidis. "Computational Physics, Molecular Dynamics Simulations". Lecture notes.

[13] G. E. Uhlenbeck and L. S. Ornstein. "On the Theory of the Brownian Motion". In: *Physical Review* 36.5 (1930), 823–841. DOI: 10.1103/physrev.36.823.

[14] Brooke E. Husic and Vijay S. Pande. "Markov State Models: From an Art to a Science". In: *Journal of the American Chemical Society* 140.7 (2018), pp. 2386–2396. DOI: 10.1021/jacs.7b12191.

[15] Erhan Cinlar. *Introduction to Stochastic Processes*. Prentice-Hall, 1975.

[16] Paul A. Gagniuc. *Markov chains from theory to implementation and experimentation*. Wiley, 2017, pp. 9–11.

[17] Joseph Rabinoff and Dan Margalit. *Interactive Linear Algebra*. Georgia Institute of Technology, 2018, 312–326. URL: https://textbooks.math.gatech.edu/ila/.

[18] Ch. Schütte and M. Sarich. "A critical appraisal of Markov state models". In: *The European Physical Journal Special Topics* 224.12 (2015), pp. 2445–2462. DOI: 10.1140/epjst/e2015-02421-0.

[19] Ch. Schütte. "Conformational Dynamics: Modelling, Theory, Algorithm, and Application to Biomolecules". Habilitation thesis at the Free University of Berlin. 1999.

[20] C. R. Schwantes, R. T. Mcgibbon, and V. S. Pande. "Perspective: Markov models for long-timescale biomolecular dynamics". In: *The Journal of Chemical Physics* (2014). DOI: 10.1063/1.4895044.

[21] Gregory R. Bowman. *An introduction to Markov State Models and their application to long timescale molecular simulation*. Springer, 2014.

[22] John D. Chodera et al. "Long-Time Protein Folding Dynamics from Short-Time Molecular Dynamics Simulations". In: *Multiscale Modeling and Simulation* 5.4 (2006), 1214–1226. DOI: 10.1137/06065146x.

[23] Purushottam D. Dixit. "Stationary properties of maximum-entropy random walks". In: *Physical Review E* 92.4 (2015). DOI: 10.1103/physreve.92.042149.

[24] Fabian Knoch and Thomas Speck. "Cycle representatives for the coarse-graining of systems driven into a non-equilibrium steady state". In: *New Journal of Physics* 17.11 (2015). DOI: 10.1088/1367-2630/17/11/115004.

[25] Udo Seifert. "Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem". In: *Physical Review Letters* 95.4 (2005). DOI: 10.1103/physrevlett.95.040602.

[26] J. Kurchan. "Fluctuation theorem for stochastic dynamics". In: *Journal of Physics A: Mathematical and General* 31.16 (1998). DOI: 10.1088/0305-4470/31/16/003.

[27] M. J. D. Powell. "On search directions for minimization algorithms". In: *Mathematical Programming* 4.1 (1973), 193–201. DOI: 10.1007/bf01584660.

[28] William C. Swope, Jed W. Pitera, and Frank Suits. "Describing Protein Folding Kinetics by Molecular Dynamics Simulations". In: *The Journal of Physical Chemistry B* 108.21 (2004), 6571–6581. DOI: 10.1021/jp037421y.

[29] Marcus Weber and Konstantin Fackeldey. "G-PCCA: Spectral Clustering for Non-reversible Markov Chains". Takustr. 7, 14195 Berlin, 2015.