Nonequilibrium Markov State Modeling

Theory and Application

Dissertation
zur Erlangung des Grades
„Doktor
der Naturwissenschaften“
am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität
in Mainz

JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

vorgelegt von
Fabian Knoch
geboren in Marburg

Mainz, April 2017

Tag der Prüfung: 07.11.2017

Erst der Enthusiasmus, dann der Fleiß.
— Stefan Zweig

To my wife

Abstract

Markov state modeling has been proven to be a powerful tool for understanding the long-term dynamics of molecular systems in thermal equilibrium. However, many striking properties of soft matter systems are inherently out of thermal equilibrium. Here we present a novel prescription extending the concept of Markov state modeling to systems that are driven into a nonequilibrium steady state. In more detail, we derive a systematic and dynamically consistent coarse graining approach for nonequilibrium Markov state models which dynamics break detailed balance. The coarse graining involves the identification and clustering of probability cycles, as well as renormalization of microscopic transition rates preserving the entropy production of the original Markov model. We apply our coarse graining procedure to a polymer under shear flow. Moreover, we show how to apply nonequilibrium Markov state modeling to periodically driven systems, which, in addition, allows us to predict the system's dynamics for different oscillation periods and protocols. Furthermore, we show how to construct nonequilibrium Markov state models for systems driven out of equilibrium due to general time-dependent driving protocols. We demonstrate the latter by an illustrative example: The Calix[4]arene-catenane dimer, a large organic complex, manipulated by a time-dependent force. Finally, we investigate the unfolding dynamics of two peptides biased by constant mechanical forces.

Zusammenfassung

Die Methodik der Markow-Zustandsmodellierung (engl. *Markov state modeling*) hat sich in den letzten Jahren als überaus hilfreiche Analysemethode zur Berechnung und Beschreibung der Langzeit-Dynamik von molekularen Systemen im thermischen Gleichgewicht bewährt. Viele verblüffende Phänomene der weichen Materie treten jedoch nur im Nichtgleichgewicht auf und bedürfen daher einer anderen Beschreibung. In dieser Arbeit werden Markow-Modelle zur Beschreibung eines stationären Nichtgleichgewichtszustandes aufgestellt, die die mikroskopische Eigenschaft des detaillierten Gleichgewichts nicht erhalten. Weiterhin wird eine neue Methode zur systematischen Vergröberung dieser Nichtgleichgewichts-Markow-Modelle eingeführt. Diese neue Vergröberungsmethode beinhaltet die Identifizierung und Gruppierung von Wahrscheinlichkeitszyklen, die als Konsequenz des Nichtgleichgewichtszustandes im ursprünglichen Markow-Modell auftreten, sowie die mathematische Umgewichtung der mikroskopischen Übergangsraten. Der dafür hergeleitete Umgewichtungsalgorithmus erhält dabei die Entropieproduktion des zugrunde liegenden Markow-Modells. Diese Vergröberungsstrategie wird anhand eines Beispiels, ein Polymer im Scherfluss, verdeutlicht. Anschließend wird gezeigt, wie Nichtgleichgewichts-Markow-Modelle für Systeme konstruiert werden können, die sich in einem stationären Nichtgleichgewichtszustand aufgrund eines oszillierenden äußeren Feldes befinden. Die dafür eingeführte Methode erlaubt zusätzlich eine Vorhersage der Dynamik des jeweiligen Systems für verschiedene Oszillationsperioden und -protokolle. Zusätzlich wird diese Methodik dahin gehend erweitert, dass auch Systeme, die durch explizit zeitabhängige Kräfte ins Nichtgleichgewicht getrieben werden, modelliert werden können. Zur Veranschaulichung wird das Calix[4]arene-Catenane Dimer, ein größerer organischer Komplex, betrachtet, welcher mithilfe einer zeitabhängigen Kraft auseinander gezogen wird. Zum Schluss wird die Entfaltungsdynamik zweier Peptide, die durch eine konstante externe Kraft entfaltet werden, untersucht.

Contents

List of Figures

Acronyms & Nomenclature

FPC	fuzzy partition coefficient
NE-MSM	nonequilibrium Markov state model
NESS	nonequilibrium steady state
MD	molecular dynamics
MFPT	mean first passage time
MSM	Markov state model
PCA	principal component analysis
PCCA+	robust Perron-Cluster cluster analysis
PMF	potential of mean force
TICA	time-lagged independent component analysis
TRAM	transition-based reweighting method

Ω	phase space
Γ	configuration space (reduced phase space)
$\mathbf{q}$	vector containing all particle positions
$\mathbf{p}$	vector containing all particle momenta
$\mathbf{r}$	$(\mathbf{p},\mathbf{q})$ point in phase space
ρ	phase space density
τ	lag time
μ	stationary (continuous) phase space density
$H(\mathbf{q},\mathbf{p})$	Hamiltonian
$U(\mathbf{q})$	potential energy
$k_B T$	thermal energy
γ_{fric}	friction coefficient
$\boldsymbol{\eta}(t)$	random force
$\mathscr{T}, \mathscr{W}$	time-discrete (-continuous) phase space propagator
$\boldsymbol{T}, T_j^i$	transition probability matrix, single transition probability ($S_i \to S_j$)
$\boldsymbol{W}, W_j^i$	transition rate matrix, transition rate from $S_i \to S_j$
Λ_i, λ_i	ith eigenvalue of $\mathscr{T}, \mathscr{W}$ or $\boldsymbol{T}, \boldsymbol{W}$
$\Phi_i(\mathbf{r})$	ith eigenfunction or eigenvector of $\mathscr{T}, \mathscr{W}$ or $\boldsymbol{T}, \boldsymbol{W}$
D	diffusion coefficient
S_i	ith discrete micro state
$\mathbf{R}_i$	configuration space position of ith centroid

$\boldsymbol{C}, c_j^i$	count matrix, transition count ($S_i \to S_j$)
p_i	microscopic probability of state S_i
π_i	steady-state probability of state S_i
$\boldsymbol{M}$, $M^{(i)}$	eigenvector matrix, ith eigenvector (PCA component)
$\boldsymbol{\Phi}, \Phi_j^i$	probability flux matrix, probability flux from $S_i \to S_j$
$\boldsymbol{J}, J_j^i$	probability current matrix, probability current from $S_i \to S_j$
f	macroscopic affinity (generalized thermodynamic force)
A_j^i	microscopic affinity (total entropy) for transition $i \to j$
σ_j^i	medium entropy produced in transition $i \to j$
s_α	entropy production rate of cycle α
$\langle \dot{S}_{\mathrm{tot}} \rangle$	mean entropy production rate
φ_α	cycle weight of cycle α
$\mathscr{G}(V, E, T)$	graph with states V, edges E and edge weights given by $\boldsymbol{T}$
$\chi_\alpha^i, \chi_{j,\alpha}^i$	indicator and passage function
A_l	cycle affinity of the lth cycle
$c_\alpha^{(i)}$	ith component of the center of cycle α
$d_\alpha^{(i)}$	ith component of the diameter of cycle α
$\mathtt{S}_l$	entropy production rate of the lth cycle community
Q	modularity function
T	temperature, protocol period
$t^{(i)}$	ith TICA component
R_{ee}	end-to-end distance
F_i	free energy of state S_i

Introduction

Since the dawn of computational physics in the late 1940s, computers have greatly extended the range of solvable problems. In particular, many-body systems – lying at the heart of statistical physics – can be efficiently treated through numerical integration. Although the development of computational physics was originally driven by military application, its first milestone, the Monte Carlo simulation technique soon triggered its application to general fields in physics and chemistry. Of particular importance is the Metropolis-Hastings algorithm [88], which is up to now the most prominent and influential Monte Carlo simulation technique. Besides its convenient implementation, its effectiveness stems from the fact that it randomly samples from the (known) Boltzmann distribution which connects ensemble probabilities with the Hamiltonian of the system. The advantage of randomly sampling the ensemble probability distribution is also the drawback of Monte Carlo simulations because, strictly speaking, one cannot infer dynamical information from a sequence of randomly sampled configurations*.

With the help of Monte Carlo simulations, classical and later quantum statistical mechanics became an intriguing research field, allowing the comparison of analytically solvable systems or analytical approximations with numerical solutions. One highlight, illustrating the impact of Monte Carlo simulations, is the understanding and prediction of phase diagrams [18, 80].

In the late 1950s, the first molecular dynamics (MD) simulations were conducted [109, 3], in which Newton's equations of motion are integrated for N particle systems. Although from a theoretical perspective the computation seems straightforward, many different numerical approximation schemes have been proposed (and sometimes reinvented), of which the Velocity Verlet algorithm [130] is the most popular example. In addition to approximating Newton's equations of motion, one has to introduce additional constraints to accurately account for thermodynamic state variables such as temperature or pressure [50].

Even if computationally more demanding and more complex to implement, MD simulations offer full insight into the dynamics of a system which allows, for example, the computation of microscopic rates.

*Although for certain sampling restrictions (local Monte Carlo moves), a physical dynamics can be approximated.

With both tools at hand and the so far steadily increasing computer power, the field of computational statistical physics grew very quickly and started to inspire related research fields. Nowadays, the interdisciplinary field of soft matter research embraces examples from physical, chemical, material and biological sciences with its focus on a variety of systems such as liquids, gels, foams, polymers, colloidal systems, granular materials and biomolecular systems. In particular, the investigation of biomolecular systems, e.g., proteins, membranes and DNA, is an increasingly active field of research, where much remains to be discovered.

During the last two decades, however, MD simulations became the method of choice for numerically investigating soft matter systems. Amongst others, this trend has been stimulated by the availability of highly optimized, easy to handle and freely accessible molecular dynamics software packages such as Gromacs [1] or LAMMPS [105] as well as the need to obtain dynamical quantities, e.g., reaction rates, diffusion coefficients, etc.. In particular, understanding the function of molecular systems, e.g., the role of a specific protein, or designing them to fulfill a specific task, makes the computation of their dynamics inevitable.

From a computational point of view, the difficulty of simulating soft matter systems arises as they are composed of unique structures ranging from a few thousands up to millions of atoms (particles) and even more electrons if a quantum mechanical treatment is necessary. Moreover, dynamically important time scales, e.g., the folding time of proteins [96, 77], range from nanoseconds up to seconds, while the time step of an all-atom classical MD simulation is bound to 2 fs, due to the very fast molecular vibrations of hydrogen.

An overview of different time and length scales, which are numerically accessible for typical soft matter systems, are illustrated in Figure 1. When electronic degrees of freedom are of interest sub nanometer length scales and femtosecond time scales are of importance. In all atomistic simulations electronic degrees of freedom, however, are neglected, allowing system sizes in the nanometer regime and typically nanonsecond to microsecond time scales. In colloidal systems a colloid is represented by a single particle with diameter ranging from tens of nanometers to a few micrometers, and effective interactions. Colloidal systems in computer simulations typically cover micrometers up to a few millimeters and time scales up to seconds. Finally, the simulation of macroscopic time and length scales is accomplished by employing a continuum mechanical description.

Unfortunately, many intriguing physical properties of soft matter systems arise from collective interactions covering many orders of magnitude in both length and time scales. A prominent example is the folding of a protein, which can take seconds, even though the protein is only a few nanometers in size. To overcome this problem, structure-based coarse graining techniques [92, 133, 143] focus on reducing the number of degrees of freedom to a fewer number of relevant ones by grouping small molecules (usually solvent molecules) or chemical groups or residues into effective particles. With less degrees of freedom, either the duration of the simulation or the effective system size can be increased. However, an intrinsic problem of structure-based coarse graining is, besides identifying the relevant degrees of freedom, that

the internal dynamics of the coarse-grained system is nonuniformly accelerated compared to the original dynamics. Loosely phrased, this acceleration stems from the fact that the effective (coarse-grained) potential energy landscape is much smoother than the original landscape, which reduces the amount of time the system spends in the vicinity of local energy minima.

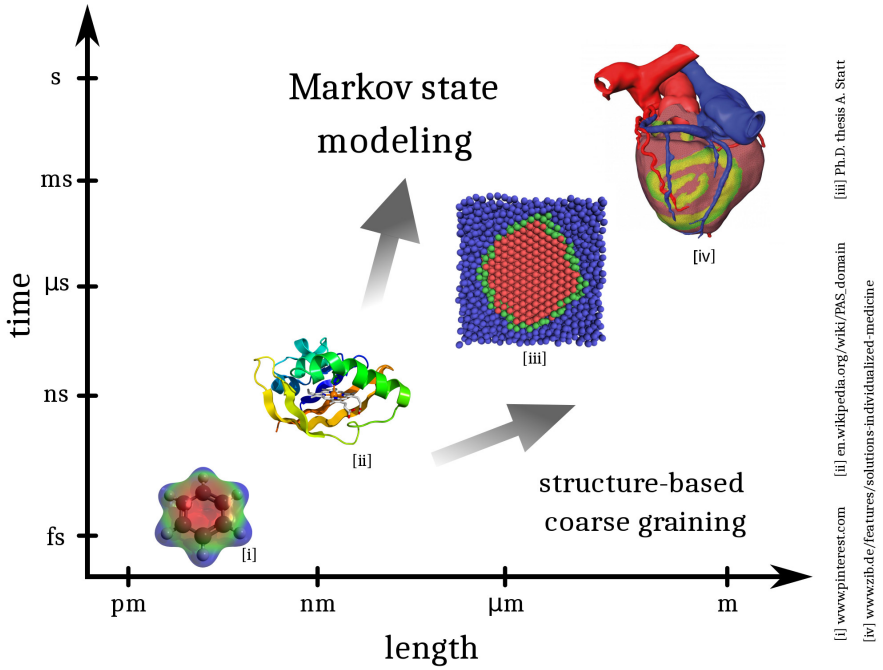

Figure 1 – **Overview of time and length scales encountered for different modeling resolutions.** The depicted systems illustrate the different levels of modeling resolutions in computer simulations. From bottom to top: The electron density of benzene (quantum mechanical modeling), an atomistic structure of a protein domain (classical all-atom molecular dynamics), colloidal particles forming a crystal nucleus (Brownian dynamics), and a electrophysiological heart model (continuum-mechanical modeling). The gray arrows represent two different strategies how to systematically coarse grain the occurring length and time scales. While the structure-based coarse graining approach is mainly developed to bridge length scales, Markov state modeling is employed to predict slow time scales.

Up to today, it is an open question how to systematically minimize this effect or how to generally relate dynamically quantities, e.g., diffusion coefficients, obtained from coarse-grained simulations to their correct values. One should note that a rigorous approach to dynamically consistent coarse graining exists for some time known as projection operator technique [53, 153]. Here, the information of the reduced variables enter in form of a memory kernel, which has to be obtained by solving a nontrivial integro-differential equation.

Quite a different route is pursued in Markov state modeling [107, 97, 21], a dynamically consistent coarse graining approach developed to bridge the long time scales involved in many biomolecular systems, e.g., the folding of proteins from an initially disordered coil to the native state. Markov state models (MSMs) are constructed for a discretized phase space

representation which discrete volumes (micro states) are further coarse grained (clustered) into a few mesoscopic states that are kinetically distinct, i.e., they correspond to basins (phase space regions) that are separated by free energy barriers with a time scale separation between fast intra-basin transitions and slow inter-basin transitions. In the case of a complete time scale separation this can be exploited to endow the mesoscopic states with a Markovian dynamics for the slow transitions while assuming quasiequilibrium for the fast transitions. Notable examples illustrating the predictive power of Markov state modeling involve the folding of small proteins [97, 141, 20, 84] or the accurate estimation of ligand-binding kinetics [104].

From a computational perspective, the key advantage of Markov state modeling is that MSMs can be constructed from many short trajectories, inferring transition rates of processes that are much longer than individual trajectories. Thus Markov state modeling is a coarse graining technique mainly designed to estimate the slow modes of the system as opposed to structure-based coarse graining which is typically employed to explore larger system sizes.

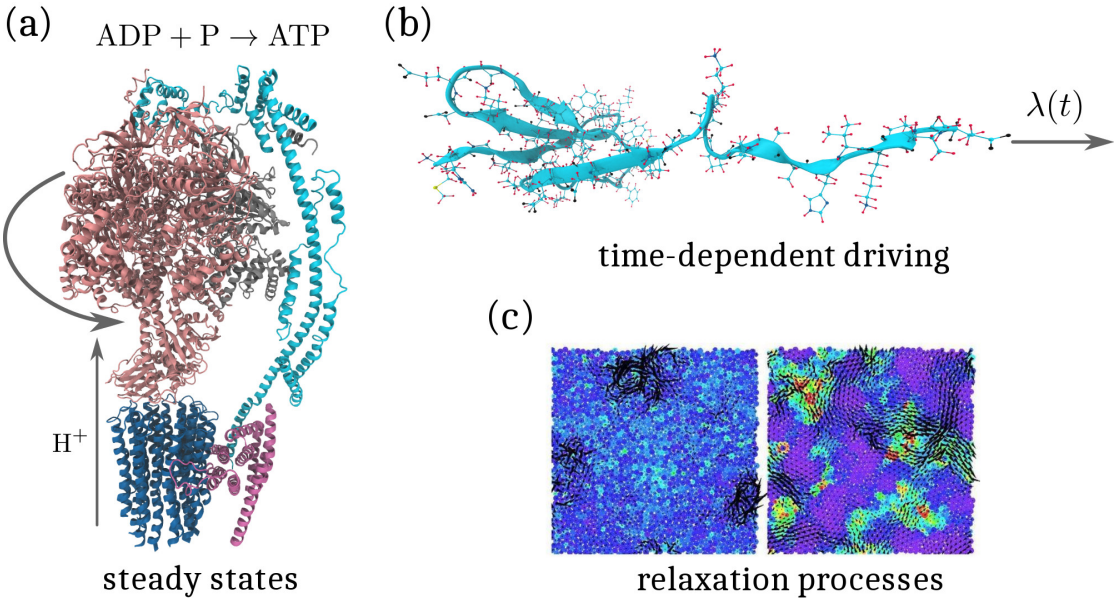

Figure 2 – **Overview of different types of nonequilibrium processes.** In nonequilibrium steady states a system is driven out of equilibrium through a time-independent driving mechanism, while for time-dependent driving the mechanism explicitly depends on time. Lastly, in nonequilibrium relaxation processes a system has been prepared in a nonequilibrium state and tries to relax back toward thermal equilibrium. Each of these types of nonequilibrium processes is represented by an exemplary system. (a) F-ATPase enzyme synthesizing ATP driven by a time-independent proton current. (b) Cold shock protein pulled by a time-dependent protocol $\lambda(t)$. (c) Nonaffine rearrangements after shear deformation of a melt (left) and a glassy state (right). While the melt relaxes quickly back toward thermal equilibrium, the particle motion for the glassy state slows down enormously and becomes strongly cooperative, illustrated by large localized rearrangements (left) and small spatially correlated rearrangements (right), respectively.[⋆].

[⋆]The subfigure has been taken from http://www.polyphys.mat.ethz.ch/research/topics/glass-transition.html.

So far, the vast majority of computer simulations of soft matter systems are performed at thermal equilibrium. In reality, however, many systems, especially biomolecular systems in their native environment, do not function at thermal equilibrium but are driven out of equilibrium. Generally, out of equilibrium or nonequilibrium processes are classified by three categories, an example of each is illustrated in Figure 2.

The first category includes systems which are driven into a nonequilibrium steady state (NESS), that is, their governing dynamics do not depend on time. The driving mechanisms responsible for maintaining a NESS consume some form of energy, e.g., mechanical or chemical energy, which has to be constantly supplied from outside the system. Possible driving mechanisms are, for instance, nonconservative forces, hydrodynamic flows, temperature differences or chemical reactions [26, 125]. Figure 2(a) illustrates an intriguing example taken from biology; the F-ATPase enzyme synthesizing ATP (adenosine triphosphate) from ADP (adenosine diphosphate) and inorganic phosphate. The reaction is catalyzed through a constant proton flux which causes the F_1 subunit (light pink and cyan colored parts) to rotate.

The second category comprises systems brought out of thermal equilibrium by time-dependent protocols, which can be given, for instance, by explicitly time-dependent mechanical forces. The example presented in Figure 2(b) shows a protein that is pulled by an external time-dependent mechanical force, which is typically realized in single-molecule force spectroscopy [113, 93].

Finally, the last category represents nonequilibrium relaxation processes, i.e., the system has been prepared in a nonequilibrium state and tries to relax back toward equilibrium. In particular, nonequilibrium relaxation processes play an important role in understanding the structural relaxation in glasses, which is also referred to as aging process [35, 16]. In Figure 2(c), the structural relaxation of a shear deformed melt (left panel) and a glass (right panel) are visualized illustrating different relaxation behaviors.

Although a general theory covering all these different nonequilibrium phenomena is currently out of reach, recently the field of stochastic thermodynamics has emerged providing a comprehensive theoretical framework in particular for driven systems (first and second category) that are still in contact with a heat reservoir that itself remains in equilibrium [124, 125]. The aim of this thesis is to exploit this framework to extend Markov state modeling to systems driven out of thermal equilibrium.

This thesis is structured as follows: The first chapter gives a comprehensive theoretical and practical introduction to equilibrium Markov state modeling as well as a brief summary of the – for this thesis important – results of stochastic thermodynamics for discrete systems. In the second chapter, we develop a general numerical routine how to construct Markov state models for nonequilibrium steady states, which we illustrate by an intuitive example: a particle in a two-dimensional double well potential, driven by a nonconservative force. In the third chapter this newly developed routine is applied to a more complex system: a tethered polymer in shear flow. Chapter four extends the developed routine to systems driven into a nonequilibrium

steady state by time-dependent but periodic protocols, for which we give two examples. In chapter five, we construct Markov state models for general time-dependent driving protocols, allowing us to mimic force-probe experiments with experimentally accessible pulling velocities. Lastly, chapter six demonstrates how Markov state models can reveal the influence of constant external mechanical forces on the unfolding pathways of small peptides.

1 Background

We start this chapter by briefly introducing the mathematical concepts of Markov state modeling. For a more thorough discussion we refer the reader to refs. [107, 21]. In the second part, we then discuss how to practically construct equilibrium Markov state models and illustrate all taken steps by an intuitive example: an one-dimensional particle in a double-well potential. In the last part, we introduce an experimentally probed Markov model for a system driven into a nonequilibrium steady state and draw the connection to stochastic thermodynamics.

1.1 Mathematical background

Consider the phase space Ω containing all dynamical variables required to describe the instantaneous state of the system. The phase space Ω can be either continuous, e.g., $\Omega \mathrel{\widehat{=}} \mathbb{R}^d$ or discrete, e.g., $\Omega \subset \mathbb{N}$. Every point in Ω at time t is described by a $d-$dimensional vector $\mathbf{r}(\mathbf{p},\mathbf{q},t) \in \mathbb{R}^d$ with $\mathbf{q}$ denoting positions and $\mathbf{p}$ momenta of all particles that constitute the system. To distinguish a dynamical process in Ω, given by a sequence of states or **trajectory**, from an instantaneous state, we refer to trajectories as $\mathbf{x}(\mathbf{r}(t)) \equiv \mathbf{x}(t)$. A trajectory $\mathbf{x}(t)$ is a particular realization of a stochastic process for which we assume that its underlying equations of motion are modeled by one of the following dynamics.

1.1.1 Thermostated Hamiltonian dynamics

The thermostated Hamiltonian dynamics is based on Newton's (or rather Hamilton's) equations of motion. As the general goal of most MD simulations is to describe an open but conservative molecular systems embedded in a heat bath with fixed temperature[⋆], one has to introduce additional dependencies referred to as thermostats. The effective equations of

[⋆] Additionally, one can impose a constant pressure by employing barostats, e.g., the Berendsen [15] or Parrinello-Rahman [102] barostat.

motion follows as

$$\frac{\partial \mathbf{q}_k}{\partial t} = \frac{1}{m_k}\,\mathbf{p}_k \qquad \text{and} \qquad \frac{\partial \mathbf{p}_k}{\partial t} = -\nabla_k U(\mathbf{q}) - \gamma(\mathbf{q},\mathbf{p})\,\mathbf{p}_k + \mathbf{F}_k^{\text{ext}}, \tag{1.1}$$

where $\mathbf{q}_k$ is the position of the kth particle, $\mathbf{p}_k$ its momentum and m_k its mass. Molecular interactions are governed by the gradient of the potential energy, which solely depends on the particle positions $\mathbf{q}$. Coupling the system with to a heat bath is typically realized by employing additional functions, here represented by $\gamma(\mathbf{q},\mathbf{p})\,\mathbf{p}_k$, which mimic the effect of a thermostat on the system. These thermostating functions are normally chosen such that the solution of eq. (1.1) conserves either kinetic or total energy. Commonly employed thermostats for molecular dynamics simulations are, amongst others, the Nosé-Hoover thermostat [98, 61], the Andersen thermostat [9] or velocity-rescaling [25]. The term $\mathbf{F}_k^{\text{ext}}$ represents, for example, external fields or nonconservative forces (forces that cannot be represented by the gradient of the potential energy).

1.1.2 Langevin dynamics

As a second type of dynamics, we consider Langevin dynamics, originally developed by Paul Langevin for the stochastic treatment of open molecular systems.

The **Langevin equations**

$$\frac{\partial \mathbf{q}_k}{\partial t} = \frac{1}{m_k}\,\mathbf{p}_k \qquad \text{and} \qquad \frac{\partial \mathbf{p}_k}{\partial t} = -\nabla_k U(\mathbf{q}) - \gamma_{\text{fric}}\,\mathbf{p}_k + \boldsymbol{\eta}_k(t) \tag{1.2}$$

can be derived from eq. (1.1) for a constant friction coefficient, i.e., $\gamma(\mathbf{q},\mathbf{p}) \equiv \gamma_{\text{fric}}$ and random force $\mathbf{F}_k^{\text{ext}}(t) \equiv \boldsymbol{\eta}_k(t)$ with zero mean $\langle \boldsymbol{\eta}_k(t) \rangle = 0$ and correlations $\langle \boldsymbol{\eta}_k^{\alpha}(t)\,\boldsymbol{\eta}_l^{\beta}(t') \rangle = \sqrt{2k_B T \gamma_{\text{fric}}}\,\delta(t-t')\,\delta_{kl}\,\delta^{\alpha\beta}$. The upper indices indicate spatial components. Here k_B denotes Boltzmann's constant and T the temperature of the surrounding heat bath. Often, one solves the Langevin equations only for the molecular system of interest and treats interactions with solvent molecules implicitly by the random force.

Furthermore, as in the case of thermostated Hamiltonian dynamics, additional nonconservative forces or external fields can be added to eq. (1.2).

In the limit of vanishing inertia ($\dot{\mathbf{p}}_k \approx 0$), eq. (1.2) takes a simpler form known as **overdamped Langevin equation** given by

$$\frac{\partial \mathbf{q}_k}{\partial t} = -\nabla_k U(\mathbf{q})/\gamma_{\text{fric}} + \mathbf{f}_k(\mathbf{q}) + \boldsymbol{\eta}_k(t). \tag{1.3}$$

The correlation of the random force $\langle \boldsymbol{\eta}_k^{\alpha}(t)\,\boldsymbol{\eta}_l^{\beta}(t') \rangle = \sqrt{2D}\,\delta(t-t')\,\delta_{kl}\,\delta^{\alpha\beta}$ can be expressed more conveniently by making use of the Einstein relation ($D = k_B T/\gamma_{\text{fric}}$), which connects the diffusion coefficient D with thermal energy and the friction coefficient. Moreover, in eq. (1.3) we explicitly include an additional function $\mathbf{f}(\mathbf{q})$ accounting for nonconservative forces or

hydrodynamic flow fields.

1.1.3 Detailed requirements

Whatever type of dynamics is chosen to model molecular systems, their associated trajectories $\mathbf{x}(t)$ are required to satisfy to following properties:

1. $\mathbf{x}(t)$ is a Markov process. The time evolution of $\mathbf{x}(t')$, i.e., $\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}\big|_{t=t'}$ (time continuous) or $(\mathbf{x}(t'+\tau)-\mathbf{x}(t'))/\tau$ (discrete), is solely determined by the position $\mathbf{r}(t')$ in phase space at time t' and does not depend on the position at previous times. Thus we can describe the time evolution between states by transition probability densities

$$p(\mathbf{r}(t),\mathbf{r}'(t+\tau)) = \mathbb{P}[\mathbf{x}(t+\tau)\in\mathbf{r}'+\mathrm{d}\mathbf{r}' \mid \mathbf{x}(t)=\mathbf{r}] \quad \forall \mathbf{r},\mathbf{r}'\in\Omega,\tau\in\mathbb{R}_{0+} \tag{1.4}$$

   expressing the conditional probability that the system evolves from state $\mathbf{r}$ at time t to an infinitesimal region $\mathrm{d}\mathbf{r}'$ around $\mathbf{r}'$ at time $t+\tau$.

2. $\mathbf{x}(t)$ is time homogeneous, i.e., the transition probability densities do not depend on the absolute time but only on the time interval τ, i.e.,

$$p(\mathbf{r}(t),\mathbf{r}'(t+\tau)) = p(\mathbf{r}(t'),\mathbf{r}'(t'+\tau)) \equiv p(\mathbf{r},\mathbf{r}';\tau) \quad \forall\, t,t'\in\mathbb{R}_{0+}. \tag{1.5}$$

   Obtaining transition probabilities between two finite regions G_0 and G_1 in Ω is achieved by integrating the transition probability densities over both regions

$$p(G_0,G_1;\tau) = \int_{G_0}\mathrm{d}\mathbf{r}\int_{G_1}\mathrm{d}\mathbf{r}'\; p(\mathbf{r},\mathbf{r}';\tau). \tag{1.6}$$

3. $\mathbf{x}(t)$ is **ergodic**, i.e., the phase space Ω cannot be separated in dynamically disjunct subspaces. The probability to reach any point $\mathbf{r}'$ in Ω from any initial point $\mathbf{r}$ within finite time is nonzero, i.e.,

$$p(\mathbf{r},\mathbf{r}';\tau) > 0 \;\; \forall \mathbf{r},\mathbf{r}'\in\Omega,\tau\in\mathbb{R}_{+}. \tag{1.7}$$

4. For $t\to\infty$, every state $\mathbf{r}$ is infinitely often visited with $\mu(\mathbf{r})$ denoting its **unique stationary density** (invariant measure) requiring that $\int_\Omega \mathrm{d}\mathbf{r}\;\mu(\mathbf{r})=1$. Defining a function $f:\Omega\to\mathbb{R}^d$, its running average is equivalent to its average on Ω weighted with $\mu(\mathbf{r})$, that is,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\; f(\mathbf{x}(t)) = \int_\Omega \mathrm{d}\mathbf{r}\; f(\mathbf{r})\mu(\mathbf{r}). \tag{1.8}$$

1.1.4 Propagator approach

Instead of describing a single trajectory $\mathbf{x}(t)$ in Ω, we now give an ensemble description. Assume multiple copies of a system are prepared in a given initial state in Ω and do not interact with each other. At time t, each trajectory is a realization of the same dynamics, while they collectively give an ensemble of visited states for every point in time. Knowing the ensemble allows the definition of a weighted time-dependent probability density function

$$\rho \,:\, \mathbf{r}(t) \in \Omega \rightarrow \rho(\mathbf{r},t) \in \mathbb{R}_{0+} \quad \text{with } 1 = \int_\Omega \mathrm{d}\mathbf{r}\, \rho(\mathbf{r},t). \tag{1.9}$$

To evolve $\rho(\mathbf{r},t)$ in time we define the **propagator** $\mathscr{T}(\tau)$ as follows:

$$\rho(\mathbf{r},t+\tau) = \mathscr{T}(\tau) \circ \rho(\mathbf{r},t) = \int_\Omega \mathrm{d}\mathbf{r}'\; p(\mathbf{r}',\mathbf{r},t)\rho(\mathbf{r}',t) \tag{1.10}$$

with $p(\mathbf{r},\mathbf{r}',t)$ being the transition probability densities. For the purpose of readability eq. (1.10) is abbreviated as

$$\rho(t+\tau) = \mathscr{T}(\tau)\rho(t). \tag{1.11}$$

When applying the propagator for infinitely long τ (lag time) the probability density converges to the stationary density $\mu(\mathbf{r})$, i.e.,

$$\lim_{\tau\rightarrow\infty} \mathscr{T}(\tau)\rho(t) = \mu. \tag{1.12}$$

It is important to note that $\mathscr{T}(\tau)$ fulfills the **Chapman-Kolmogorov equation**

$$\rho(t+k\tau) = \mathscr{T}(k\tau)\rho(t) = \left[\mathscr{T}(\tau)\right]^k \rho(t),\; k\in\mathbb{N}, \tag{1.13}$$

stating that it is equivalent to evolve the system in time once with lag time $k\tau$ or to evolve it $k$times successively with lag time τ.

Before discussing a second approach, it is, however, instructive to first draw the connection to equilibrium statistical mechanics.

1.1.5 Equilibrium statistical mechanics

A Markov process $\mathbf{x}(t)$ associated with a dynamics in thermal equilibrium is reversible, i.e., the transition probability densities [eq. (1.7)] fulfill the **detailed balance** condition

$$\mu(\mathbf{r})\,p(\mathbf{r},\mathbf{r}';\tau) = \mu(\mathbf{r}')\,p(\mathbf{r}',\mathbf{r};\tau), \tag{1.14}$$

which expresses that the fraction of the system transported from $\mathbf{r}$ to $\mathbf{r}'$ in time τ is the same as the fraction transported from $\mathbf{r}'$ to $\mathbf{r}$. In other words, there is no net probability flux between both states, or on average the systems traverses from $\mathbf{r}$ to $\mathbf{r}'$ the same number of times as from

$\mathbf{r}'$ to $\mathbf{r}$. Detailed balance hence implies that it is not possible to transport any physical quantity, e.g., charge, particles, etc.. This critical restriction, however, allows to connect the stationary density $\mu(\mathbf{r})$ with the Hamiltonian of the system (Boltzmann distribution), given by

$$\mu(\mathbf{r}) = \frac{e^{-\beta H(\mathbf{q},\mathbf{p})}}{Z} \quad \text{with} \quad Z = \int_\Omega \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}\; e^{-\beta H(\mathbf{q},\mathbf{p})}. \tag{1.15}$$

The factor $\beta = 1/k_B T$ ensures energy normalization, while the partition function Z ensures the normalization of probability.

1.1.6 Fokker-Planck and infinitesimal generator

As discussed in sections 1.1.2, one possibility of describing the dynamics of the system is by applying the stochastic Langevin formalism. For simplicity we consider here the overdamped Langevin equation [see eq. (1.3)], although the treatment of the full Langevin dynamics or thermostated Hamiltonian dynamics can be formulated analogously.

Instead of solving the dynamics for every particle individually, we can formulate the associated Fokker-Planck equation which describes the time evolution of the phase space density $\rho(t)$, that is,

$$\frac{\partial \rho(t)}{\partial t} = \mathscr{W}\,\rho(t) \quad \text{with } \mathscr{W} \equiv \left[D\,\Delta + \Big(-\nabla U(\mathbf{q})/\gamma_{\text{fric}} + \mathbf{f} \Big) \cdot \nabla \right]^\dagger, \tag{1.16}$$

where $\mathscr{W}$ is referred to as **Fokker-Planck generator**. Here ∇ and Δ act on all particle positions $\mathbf{q}$. Similarly, the vector $\mathbf{f}(\mathbf{q},\mathbf{p})$ represents external forces acting on all particles.

The formal solution of the Fokker-Planck equation is easily obtained if the operator $\mathscr{W}$ does not explicitly depend on time, then,

$$\rho(t) = e^{\mathscr{W} t}\rho_0, \tag{1.17}$$

where $\rho_0 = \rho(t=0)$ expresses the initial density. For $t \to \infty$, this density converges to the stationary distribution $\mu(\mathbf{r})$. Moreover, if the external force $\mathbf{f}$ is conservative, i.e., can be expressed by the negative gradient of a scalar potential energy function, the stationary solution of eq. (1.17) is the Boltzmann distribution [cf. eq. (1.15)].

Alternatively to the Fokker-Planck generator, one can define the **infinitesimal generator** by taking the continuous limit of the discrete propagator $\mathscr{T}$, i.e.,

$$\mathscr{W} \equiv \lim_{\tau \to 0} \frac{1}{\tau}\Big(\mathscr{T}(\tau) - \mathbf{1} \Big), \tag{1.18}$$

where $\mathbf{1}$ denotes the identity operator. Both, the Fokker-Planck and infinitesimal generator, are considered to be equivalent, i.e., both determine the time evolution of the phase space density $\rho(t)$, cf. eq. (1.17), and therefore the stationary density μ.

Comparing the action of the propagator $\mathscr{T}(\tau)$ [see eq. (1.11)] with eq. (1.17), we identify

$$\mathscr{T}(\tau) = e^{\mathscr{W}\tau} \tag{1.19}$$

showing that the propagator $\mathscr{T}(\tau)$ is the time-discrete counterpart of the time-continuous Fokker-Planck (infinitesimal) generator $\mathscr{W}$. Because of their relation both operators share or exhibit closely related properties, which we examine in the next section.

Finally, we want to emphasize again that for systems in thermal equilibrium the following statement is equivalent:

detailed balance holds $\longleftrightarrow$ the stationary density μ is the Boltzmann distribution

1.1.7 Properties of $\mathscr{T}$ and $\mathscr{W}$

Due to the relation in eq. (1.19), both operators share the same set of eigenfunctions ϕ_i and thus

$$\mathscr{T}(\tau)\phi_i(\mathbf{r}) = \Lambda_i\phi_i(\mathbf{r}) \qquad \text{and} \qquad \mathscr{W}\phi_i(\mathbf{r}) = \lambda_i\phi_i(\mathbf{r}) \tag{1.20}$$

with their eigenvalues λ_i and Λ_i being related through $\Lambda_i(\tau) = e^{\lambda_i\tau}$. This relation stems from the fact that both operators share the same positive semi-group [cf. Chapmann-Kolmogorov eq. (1.13)], i.e.,

$$\mathscr{T}(\tau_1)\mathscr{T}(\tau_2) = \mathscr{T}(\tau_1+\tau_2). \tag{1.21}$$

For both operators, eigenvalues and eigenvectors can be complex-valued with their real part bounded by

$$\mathrm{Re}(\Lambda_i) \in (0,1] \quad \text{and} \quad \mathrm{Re}(\lambda_i) \in (-\infty, 0] \qquad \Lambda_i, \lambda_i \in \mathbb{C}. \tag{1.22}$$

From a mathematical point of view the real part of the eigenvalue spectrum of the time-discrete propagator $\mathscr{T}(\tau)$ is bounded by $\mathrm{Re}(\Lambda_i) \in [-1,1]$ [see ref. [123]]. From a physical perspective, however, $\mathrm{Re}(\Lambda_i)$ cannot be negative as otherwise the continuous generator $\mathscr{W}$ is not guaranteed to exist [89]. In any case, the absolute operator spectra can be sorted in descending order, i.e., $|\lambda_0| > |\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$, where complex-valued eigenvalues always occur in complex conjugated pairs. The eigenfunction $\phi_0(\mathbf{r})$ associated with the largest eigenvalue (also known as Perron root), that is, $\lambda_0 = 1$ and $\Lambda_0 = 0$, respectively, is real-valued and corresponds to the stationary distribution $\mu(\mathbf{r})$.

Mathematically, we consider both operators to act on the Hilbert space of square integrable functions

$$L^2 = \left\{ f : \Omega \to \mathbb{C} : \; ||f||_2^2 = \int_\Omega \mathrm{d}\mathbf{r}\, |f(\mathbf{r})|^2 < \infty \right\} \tag{1.23}$$

equipped with the standard scalar product

$$\langle f \mid g \rangle = \int_\Omega \mathrm{d}\mathbf{r}\, f^*(\mathbf{r}) g(\mathbf{r}), \tag{1.24}$$

where * denotes complex conjugation. At this stage, it is helpful to define a set of functions $\psi_i(\mathbf{r})$ that are related to the eigenfunctions $\phi_i(\mathbf{r})$ via

$$\psi_i(\mathbf{r}) \equiv \phi_i(\mathbf{r}) \cdot \mu(\mathbf{r})^{-1}. \tag{1.25}$$

In fact, the $\psi_i(\mathbf{r})$'s are eigenfunctions of the forward transfer operator [123], which definition is analogous to $\mathscr{T}$ (eq. (1.11)) but w.r.t. the stationary measure $\mu(\mathbf{r})$ – it acts on the $\mu(\mathbf{r})$-weighted probability space. Employing the definition of $\phi_i(\mathbf{r})$, we agree on the following normalizations:

$$\langle \phi_i \mid \psi_i \rangle = 1 \quad \text{and} \quad \int_\Omega \mathrm{d}\mathbf{r}\, \phi_0(\mathbf{r}) = 1. \tag{1.26}$$

Operator spectrum for reversible Markov processes

If the Markov process $\mathbf{x}(t)$ is reversible, i.e., the detailed balance condition eq. (1.14) is fulfilled, the operators $\mathscr{T}$ and $\mathscr{W}$ are self-adjoint and hence all their eigenfunctions are orthogonal $\langle \phi_i \mid \psi_j \rangle = \delta_{ij}$ and eigenvalues are real-valued. Moreover, their eigenvalues λ_i have the physical meaning of time scales, i.e., $t_i = -\lambda_i{}^\star$, which becomes clear when expanding eq. (1.17) (the expansion of eq. (1.11) is analogous);

$$\rho(\mathbf{r}, t) = e^{\mathscr{W} t} \rho_0(\mathbf{r}) = \sum_k \langle \phi_k | \rho_0 \rangle\, e^{\lambda_k t} \phi_k(\mathbf{r}). \tag{1.27}$$

For $t \to \infty$, all modes except the mode corresponding to $\lambda_0 = 0$ have decayed to zero. The closer λ_i is to zero, the slower the associated mode decays. Thus we conclude that slow processes belong to eigenvalues close to zero – or close to one for Λ_i. Exploiting the eigenvalue spectrum, the dynamics can be separated into m slow decaying modes and the remaining fast decaying modes. If the spectrum offers a large gap between λ_m and λ_{m+1}, i.e., $\lambda_m \gg \lambda_{m+1}$ the first m modes are associated with **metastable** regions in phase space. The key idea is that it takes the system a significant longer time to traverse between states of different metastable regions than between states of the same metastable region. According to eq. (1.27), we can split the generator $\mathscr{W}$, such that

$$\rho(\mathbf{r}, \tau) = e^{\mathscr{W}_{\text{slow}} \tau + \mathscr{W}_{\text{fast}} \tau} \rho_0(\mathbf{r}) = \sum_{k=0}^{m} \langle \phi_k | \rho_0 \rangle\, e^{\lambda_k \tau} \phi_k(\mathbf{r}) + e^{\mathscr{W}_{\text{fast}} \tau} \rho_0(\mathbf{r}). \tag{1.28}$$

Identifying and preserving these slow modes is the general goal of **equilibrium Markov state modeling**.

*For eigenvalues Λ_i, the ith implied time scale is given by $t_i = -\tau / \log(\Lambda_i)$

1.1.8 Discretization

Consider a crisp discretization of phase space Ω into n sets. These n sets fully partition Ω, labeled by $S = \{S_0, ..., S_n\}$ with $\cup_{i=0}^{n} S_i = \Omega$, and do not overlap, i.e., $S_i \cap S_j = \emptyset \ \forall i \neq j$. We further define the membership functions

$$\nu_i(\mathbf{r}) = \begin{cases} 1, & \mathbf{r} \in S_i \\ 0, & \mathbf{r} \notin S_i \end{cases}. \tag{1.29}$$

which assign every point $\mathbf{r}$ in Ω uniquely to one of the S_i's.

Having established a phase space discretization, the stationary probability distribution π_i of set i is computed by integrating the stationary probability density, i.e.,

$$\pi_i = \int_\Omega \mathrm{d}\mathbf{r}\, \nu_i(\mathbf{r})\mu(\mathbf{r}) \qquad \text{, while} \qquad \mu_i(\mathbf{r}) = \begin{cases} \mu_i(\mathbf{r})/\pi_i, & \mathbf{r} \in S_i \\ 0, & \mathbf{r} \notin S_i \end{cases} \tag{1.30}$$

denotes the local stationary distribution restricted to set i. The important point here is that both π_i and $\mu_i(\mathbf{r})$ are local properties not requiring information about the full phase space Ω. To describe the transition probabilities between sets i and j, the propagator is projected onto the membership functions yielding discretized transition probabilities

$$T_{ij}(\tau) = \frac{\langle \nu_j \mid (\mathscr{T}(\tau) \circ \nu_i) \rangle}{\langle \nu_i \mid \nu_i \rangle}. \tag{1.31}$$

Every T_{ij} can be interpreted as the conditional probability that the systems jumps from state S_i to state S_j within time τ, that is,

$$T_{ij}(\tau) = \mathbb{P}[\mathbf{x}(t+\tau) \in S_j \mid \mathbf{x}(t) \in S_i]. \tag{1.32}$$

The entity of all transition probabilities forms the **transition probability matrix** $\boldsymbol{T}(\tau) \in \mathbb{R}_{0+}^{n \times n}$ (shortened from now on as transition matrix), which is row-stochastic, i.e., $\sum_j T_{ij} = 1$. The discrete analog of eq. (1.13) is therefore given by the matrix vector product

$$p^{\mathrm{T}}(t + k\tau) = p^{\mathrm{T}}(t)\boldsymbol{T}^k(\tau) \quad k \in \mathbb{N}. \tag{1.33}$$

All spectral properties of $\mathscr{T}$ hold for its discrete counterpart $\boldsymbol{T}$ too. For example, the stationary solution π is the left eigenvector of $\boldsymbol{T}$ belonging to $\Lambda_0 = 1$, i.e.,

$$\pi^{\mathrm{T}} = \pi^{\mathrm{T}}\boldsymbol{T}(\tau). \tag{1.34}$$

The discretization of the continuous generator $\mathscr{W}$ is analogous to eq. (1.31) with $\boldsymbol{W} \in \mathbb{R}^{n \times n}$ denoting the row-stochastic transition rate matrix. All off-diagonal elements are positive, i.e., $w_{ij} \geq 0$ for $i \neq j$ and represent transition rates from sets S_i to S_j, whereas the negative diagonal elements $w_{ii} = -\sum_{j \neq i} w_{ij}$ ensure probability conservation. The time evolution of

probabilities employing the transition rate matrix, given by

$$\frac{\partial p^{\mathrm{T}}(t)}{\partial t} = p^{\mathrm{T}}(t)\boldsymbol{W} \qquad \text{with} \qquad 0 = \pi^{\mathrm{T}}\boldsymbol{W}, \tag{1.35}$$

is known as the **master equation**. Hence the relation eq. (1.19) that connects both continuous operators holds true for their discrete counterparts $\boldsymbol{T}(\tau) = \exp(\boldsymbol{W}\tau)$ too.

Note that some authors, see for instance refs. [31, 23], define the transition matrix as column stochastic. Then, the T_{ij}'s represent transition probabilities from states (from now on referred to as states) S_j to S_i with normalization $\sum_i T_{ij} = 1$. The advantage of the column-stochastic representation is that eq. (1.33) is expressed by a matrix "column vector" product which is the standard implementation in most numerical libraries. Adopting this notation, eq. (1.33) becomes $p(t + k\tau) = \boldsymbol{T}^k(\tau)p(t)$ with π being a right (instead of left) eigenvector of $\boldsymbol{T}$. To avoid confusion transition probabilities/rates are from now on labeled as T^i_j / W^i_j for transitions $i \to j$.

Our task is now to approximate the transition matrix $\boldsymbol{T}$ from molecular dynamics simulations.

1.2 Constructing Markov state models in practice

Before discussing the technical details of how to best construct Markov state models (MSMs) from finite trajectories, we comment on the Markov assumption of the approximated discrete dynamics. In fact, although the true continuous dynamics is Markovian by construction, the approximated dynamics [see eq. (1.33)] is not Markovian on all time scales. The reason for this lies in the fact that information is lost when discretizing phase space. The amount of missing information can be associated with the production of entropy, which manifests itself by causing memory effects in the approximated dynamics. This non-Markovianity occurs even more strongly when, instead of approximating the full phase space, only a reduced phase space representation is discretized.

Indeed, Markov models typically project out all momenta and solvent coordinates, and moreover employ only a subset of distance coordinates, e.g., distances or angles between heavy atoms or alpha carbon atoms. Losing Markovianity and hence the occurrence of memory effects is a known issue when coarse graining degrees of freedom, which can be, for instance, tackled by introducing memory kernels as described in the Mori-Zwanzig formalism [59]. However, in Markov state modeling one assumes a clear separation between rapidly decaying time scales that corresponds to a strongly localized dynamics, e.g., bond vibrations, and slow decaying time scales that are associated with global motion, e.g., large conformational changes. The quickly decaying modes are assumed to be spatially localized within the discretized states S_i with associated time scales much faster than the lag time τ of the Markov state model. Markovianity is recovered when considering only the slow decaying time scales, whose associated collective motion covers many discrete states. Therefore, we understand

Markov models as models which approximate the long-time dynamics. How to test whether a given transition matrix $\boldsymbol{T}(\tau)$ is a good approximation of the long-term dynamics we discuss at the end of this section.

1.2.1 Phase space discretization

For constructing MSMs from molecular dynamics simulations, one assumes that the phase space has been sufficiently explored by either a few long trajectories or many short trajectories. Moreover, we assume that a "reasonable" phase space projection exists, which is referred to as configuration space Γ, i.e., $\Gamma \subset \Omega$. Depending on the dimensionality, Γ can be either discretized by equidistantly positioned bins or, which is commonly done for more than two dimensions, by automated clustering algorithms, e.g., k-means, k-centers, etc. [2].

Here we only give a brief summary of the k-means algorithm as we exclusively employ it throughout this thesis. The input of the k-means algorithm is given by the set of trajectories in Γ and an integer k indicating the number of cells that discretize the configuration space. Typically, k is of the order $10^2 - 10^3$ which is why the discretization is often referred to as fine-graining. k-means randomly places k points or **centroids** $\mathbf{R}_k$ in Γ and computes the residual

$$\mathrm{res}(k) = \sum_{i=0}^{N} \min_k \left|\left|\mathbf{r}_i - \mathbf{R}_k\right|\right|_2 \tag{1.36}$$

given by the Euclidean distance between all positions and their closest centroid. Next, new $\mathbf{R}_k$ are determined by computing the mean of all positions that were assigned to the kth centroid. All centroid positions are iteratively optimized until the residual has been converged. Knowing all centroid coordinates $\mathbf{R}_k$, the configuration space is fully discretized by assigning every point in Γ to the closest centroid $\mathbf{R}_k$. Note that this type of distance-dependent space discretizations is also referred to as Voronoi tessellation. Adopting the nomenclature introduced in the previous section, every $\mathbf{R}_k$ uniquely defines a (micro) state S_k.

1.2.2 Approximating T from molecular dynamics simulations

To approximate the transition matrix $\boldsymbol{T}$ from MD simulations, one maps all trajectories – here we assume only one trajectory – containing N frames onto the set of states, i.e.,

$$[\mathbf{r}(0), \mathbf{r}(\Delta t), \mathbf{r}(2\Delta t), ..., \mathbf{r}((N-1)\Delta t)] \quad \longmapsto \quad [s(0), s(\Delta t), s(2\Delta t), ..., s((N-1)\Delta t)], \tag{1.37}$$

where Δt is the time interval for which a frame was stored and $s(t)$ the corresponding discrete state at time t. Thus all dynamical information are effectively stored in a simple sequence. Figure 1.1 illustrates the mapping of a continuous trajectory onto a discrete sequence of states.

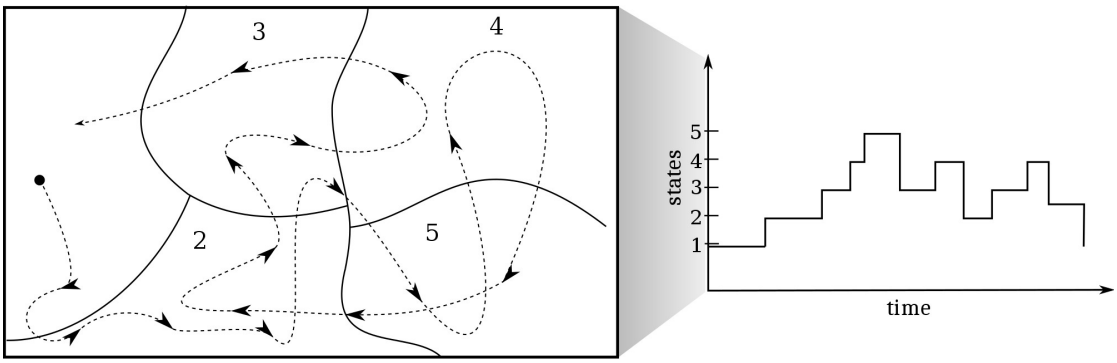

Figure 1.1 – **Phase space discretization**. The space-continuous trajectory is approximated by a Markov jump process.

To extract dynamical information from the sequence of states, we count all observed transitions $c_j^i(\tau)$ for lag time $\tau = l\Delta t$ with $l \in \mathbb{N}_+$, that is,

$$c_j^i(\tau) = c_j^i(l\Delta t) = \sum_{k=1}^{N-l} \nu_i(\mathbf{r}(k\Delta t))\, \nu_j(\mathbf{r}((k+l)\Delta t)). \tag{1.38}$$

If the **count matrix** $\boldsymbol{C}$ is associated with an infinitely long trajectory, the transition probabilities are given by the trivial estimator

$$\hat{T}_j^i = \frac{c_j^i}{\sum_i c_j^i}, \tag{1.39}$$

expressing the fraction of times the transition $i \to j$ is observed relative to all transitions starting in i. Unfortunately, if the number of observed counts are finite, one cannot uniquely determine the underlying transition matrix $\boldsymbol{T}(\tau)$. However, we can ask for the probability to observe the count matrix $\boldsymbol{C}$ given the true transition matrix $\boldsymbol{T}(\tau)$. Mathematically, this conditional probability reads

$$p(\boldsymbol{C} \mid \boldsymbol{T}) = \prod_{i,j} \left(T_j^i\right)^{c_j^i}, \tag{1.40}$$

for which one assumes that the c_j^i's are statistically uncorrelated. Employing Bayes' theorem, eq. (1.40) can be reversed, i.e.,

$$p(\boldsymbol{T} \mid \boldsymbol{C}) \propto p(\boldsymbol{T})\, p(\boldsymbol{C} \mid \boldsymbol{T}), \tag{1.41}$$

where $p(\boldsymbol{T} \mid \boldsymbol{C})$ is the posterior probability and $p(\boldsymbol{T})$ the prior probability of transition matrices after and before knowing $\boldsymbol{C}$, respectively. When no prior information are known, the prior distribution is assumed to be flat, i.e., $p(\boldsymbol{T}) \propto 1$. Our task is to find the most probable transition matrix given the count matrix $\boldsymbol{C}$, i.e., $\hat{\boldsymbol{T}} = \operatorname{argmax} p(\boldsymbol{T} \mid \boldsymbol{C})$. In statistical science this problem is referred to as finding the **maximum of the likelihood function** $p(\boldsymbol{T} \mid \boldsymbol{C})$. Indeed, one can

show [see ref. [107]] that the estimator defined for infinite counts [eq. (1.39)] maximizes the likelihood function when no further restrictions are given. However, the ergodicity condition [cf. eq. (1.7)] requires that if $T_j^i > 0$, then $T_i^j > 0$ for all $i \neq j$. This condition becomes even more restrictive when the dynamics is reversible as $\hat{\boldsymbol{T}}$ is required to obey the detailed balance condition, i.e., $T_j^i \pi_i = T_j^i \pi_j$. Unfortunately, due to the lack of an analytical solution incorporating these restrictions, a numerical solution is required. Generally, two types of approaches are employed. The first is to find $\hat{\boldsymbol{T}}$ by maximizing the logarithmic likelihood function

$$\hat{\boldsymbol{T}} = \operatorname{argmax} \log\Big(p(\boldsymbol{T} \mid \boldsymbol{C})\Big) \;=\; \operatorname{argmax} \Big(\sum_{ij} c_j^i \log\big(T_j^i\big)\Big), \tag{1.42}$$

which is commonly accomplished by employing a fixed-point iteration scheme [107] or Lagrange multipliers [136, 135].

The second approach, known as **Bayesian inference**, does not search for the maximum of the likelihood function but rather samples the distribution $p(\boldsymbol{T} \mid \boldsymbol{C})$ [134]. The sampling typically follows a Monte-Carlo sampling technique, where new transitions probabilities T_j^i are proposed and accepted or rejected according to a specific weight function. Thus, instead of returning a single transition matrix, the Bayesian approach returns an ensemble of transitions matrices. The advantage of the Bayesian approach lies in the fact that it allows a more accurate estimation of error intervals. For example, the stationary distribution given by the eigenvector ϕ_0, can be computed for the full ensemble of transition matrices, returning an ensemble of stationary distributions. An estimation of the stationary distribution is then determined by the mean of the ensemble, while the standard error is determined using its variance. However, having an ensemble of transition matrices allows the computation of general confidence intervals [134].

1.2.3 Testing for Markovianity

Lag time analysis

A popular approach to test the quality of an estimated transition matrix $\boldsymbol{T}(\tau)$, originally introduced in ref. [132], is referred to as lag time analysis.

Any constructed MSM (for a given lag time τ_0) is supposed to be a good approximation of the dynamics for lag times $\tau \geq \tau_0$, implying that the eigenvalues $\Lambda_i(k\tau_0)$ are well approximated by $[\Lambda_i(\tau_0)]^k$. However, instead of eigenvalues, one often refers to implied time scales [see section 1.1.7] as the implied time scales t_i are supposed to be approximately constant over the range of $\tau = k\tau_0$,

$$t_i(k\tau_0) = -\frac{\tau_0}{\log\left(\Lambda_i(\tau_0)\right)} \approx t_i \qquad \text{for } k = 1, 2, \ldots. \tag{1.43}$$

The lag time analysis suggests to test for the convergence of the slowest implied time scale. The smallest possible lag time τ_0 at which the slowest implied time scale is approximately constant

then serves as the lag time for which the final MSM is constructed. Although often empirically observed [30], one should note that a constant implied time scale does not strictly guarantee Markovianity of the slow dynamics. The reverse, however, holds true, i.e., Markovianity implies that the implied time scales do not depend on the lag time [131].

Chapman-Kolmogorov test

To test for Markovianity the Chapman-Kolmogorov test checks whether the approximation

$$\left[\boldsymbol{T}(\tau)\right]^k \approx \boldsymbol{T}(k\tau) \tag{1.44}$$

holds within statistical uncertainty. For large transition matrices it can be a cumbersome task to compare all elements individually, which moreover can have large uncertainties. Therefore, for the comparison one typically chooses a set of states that is in a sense essential for the dynamics, e.g., a metastable set. The comparison then should be done for as many values of k as allowed by the data set. Here we briefly review an algorithm described in ref. [21].

We start by defining a set of states, say A, for which we want to check eq. (1.44). Its corresponding stationary distribution (following the estimated transition matrix $\boldsymbol{T}(\tau_0)$) is given by

$$\omega_i^A = \begin{cases} \frac{\pi_i}{\sum_{j\in A}\pi_j} & i \in A \\ 0 & i \neq A \end{cases}. \tag{1.45}$$

Next, we define a trajectory-based time-dependent probability

$$p_{\mathrm{MD}}(A, A; k\tau) = \sum_{i\in A} \omega_i^A \left(\frac{\sum_{j\in A}\ c_j^i(k\tau)}{\sum_j\ c_j^i(k\tau)} \right) \tag{1.46}$$

indicating the occupation of set A when starting with distribution $\boldsymbol{\omega}^A$. Here the $c_j^i(k\tau)$'s are the transition counts for lag time $k\tau$.

Similar, the probability to be at set A at time $k\tau$, following the Markov model, is given by

$$p_{\mathrm{MSM}}(A, A; k\tau) = \sum_{i\in A} \left[\left(\boldsymbol{\omega}^A\right)^{\mathrm{T}}\ \boldsymbol{T}^k(\tau) \right]_i. \tag{1.47}$$

According to eq. (1.44), the Chapman-Kolmogorov property is fulfilled if both probabilities match, i.e., $p_{\mathrm{MD}}(A, A; k\tau) \approx p_{\mathrm{MSM}}(A, A; k\tau)$, within the statistical error accounted for by

$$\varepsilon_{\mathrm{MD}}(A, A; k\tau) = \sqrt{k\ \frac{p_{\mathrm{MD}}(A, A; k\tau) - [p_{\mathrm{MD}}(A, A; k\tau)]^2}{\sum_{i\in A}\sum_j c_j^i(k\tau)}}. \tag{1.48}$$

1.2.4 From transition probabilities to transition rates

So far we discussed how to best approximate the transition matrix $\boldsymbol{T}$ from sampled trajectories. However, analogous to eq. (1.19), the rate matrix $\boldsymbol{W}$ is the time-continuous counterpart of the transition matrix $\boldsymbol{T}$. Theoretically, if the relation

$$\boldsymbol{T}(\tau) = \exp(\boldsymbol{W}\tau) \tag{1.49}$$

holds true, then $\boldsymbol{W}$ can be determined by

$$\boldsymbol{W} = \frac{\boldsymbol{M}\,\log\left(\boldsymbol{D}_T\right)\,\boldsymbol{M}^{-1}}{\tau}. \tag{1.50}$$

Here we exploit the fact that both matrices share the same set of eigenvectors $\boldsymbol{M}$ allowing their diagonalization through the same linear transformation $\boldsymbol{D}_T = \boldsymbol{M}^{-1}\,\boldsymbol{T}\,\boldsymbol{M}$ and $\boldsymbol{D}_W = \boldsymbol{M}^{-1}\,\boldsymbol{W}\,\boldsymbol{M}$. In practice, however, the estimated transition matrix $\hat{\boldsymbol{T}}$ does not necessarily fulfill eq. (1.49), that is, no physical rate matrix exists with $W_j^i \geq 0 \;\; \forall i \neq j$, which is known as the **embedding problem** [64, 89]. Unfortunately, up to today no conditions for $\hat{\boldsymbol{T}}$ have been found that ensure the existence of a rate matrix, though ref. [64] provides a list of conditions for which a rate matrix cannot exist.

When a given transition matrix $\hat{\boldsymbol{T}}$ is not embeddable, it is still useful to determine an auxiliary rate matrix approximating the time-continuous dynamics. One approach described in ref. [64] is to formally invert eq. (1.49) and approximate the matrix logarithm by its series expansion, leading to

$$\hat{\boldsymbol{W}} = \frac{1}{\tau}\log\left(\hat{\boldsymbol{T}}\right) = \frac{1}{\tau}\log\left(\mathbf{1} + \hat{\boldsymbol{T}} - \mathbf{1}\right) \approx \frac{1}{\tau}\left((\hat{\boldsymbol{T}} - \mathbf{1}) - \frac{(\hat{\boldsymbol{T}} - \mathbf{1})^2}{2} + \frac{(\hat{\boldsymbol{T}} - \mathbf{1})^3}{3} - \ldots\right). \tag{1.51}$$

The expansion can be computed recursively, while we stop either if the auxiliary matrix has any negative nondiagonal entries or the change in the next order approximation is small enough. Note that the linear approximation always returns a physical rate matrix as it coincides with the linear approximation of eq. (1.49) for small τ.

Alternatively, one can estimate $\hat{\boldsymbol{W}}$ directly [87] by finding the maximum of the likelihood function, i.e.,

$$\hat{\boldsymbol{W}} = \operatorname{argmax} \log\Big(p(\boldsymbol{W} \mid \boldsymbol{C})\Big) \;=\; \operatorname{argmax} \sum_{ij} c_j^i \,\log\Big(\exp\left(\boldsymbol{W}\tau\right)_j^i\Big). \tag{1.52}$$

Here $\exp(\boldsymbol{W}\tau)_j^i$ evaluates the matrix exponential and returns its $i \to j$ element. Besides the restrictions of a rate matrix, i.e., $W_j^i > 0 \;\; \forall i \neq j$ and $\sum_i W_j^i = 0$, the maximum likelihood approach allows, analogously to its time-discrete counterpart eq. (1.39), to incorporate additional conditions such as detailed balance ($W_j^i \pi_i = W_i^j \pi_j$).

While the computation of eq. (1.52) quickly becomes more expensive than eq. (1.39), its

benefit is that many "unphysical" entries of $\hat{\boldsymbol{W}}$ become zero, yielding a more sparse and thus more realistic rate matrix. Unphysical, here, refers to transitions between states which are microscopically impossible and only occur due to the lag time τ. Suppose, for instance, the transitions $S_0 \leftrightarrow S_1 \leftrightarrow S_2$ are probed for short τ, i.e., $\{T_1^0, T_0^1, T_2^1, T_1^2\} > 0$ and $T_2^0 = T_0^2 = 0$. When estimating the transition matrix for longer τ, the same dynamics can exhibit an additional transition $S_0 \leftrightarrow S_2$, i.e., $\{T_2^0, T_0^2\} > 0$. The maximum likelihood approach for the rate matrix suppresses transition rates corresponding to this type of unphysical transitions [87]. Lastly, knowing the rate matrix allows to determine transition times by computing **mean first passage times**.

1.2.5 Computing transition times

Traditionally, reaction rates (inverse transition times) are computed by making use of the seminal result of Arrhenius [10] which was later refined by Eyring and Polanyi [47, 48], and Kramer [76]. The escape rate for leaving a free energy minimum or alternatively the transition rate from free energy basin A to free energy basin B is given (in the large friction limit) by

$$k_{A\to B} = \frac{\omega_A\,\omega_0}{2\pi\gamma_{\text{fric}}}\,\exp\big(-\beta\,\Delta F_0\big), \tag{1.53}$$

where ω_A is the angular frequency inside minimum A, ω_0 the angular frequency at the transition state separating A and B, and ΔF_0 the height of the free energy barrier [56]. The advantage of eq. (1.53) is that, if the full free energy (potential energy) surface is known, ω_A and ω_B can be estimated by the curvature of the free energy at A and at the transition point. However, in practice it is a challenging task to accurately determine ω_A and ω_0, and therefore often fixed values are assumed. The second difficulty arises because of already a small error in ΔF_0 can lead to large uncertainties in $k_{A\to B}$, due to its exponential dependence. Furthermore, when investigating systems out of thermal equilibrium, eq. (1.53) looses its validity.

In Markov state modeling, on the other hand, all dynamical information are stored in the form of a transition or rate matrix. To visualize the system's dynamics we introduce the concept of an edge-weighted **directed graph** $\mathcal{G}(V, E, W)$, where vertices V are identified with states, edges E with transitions between states and edge weights W with transition rates. In accordance with the ergodicity condition, the graph $\mathcal{G}$ is connected, that is, every vertex can be reached from every other vertex after a finite number of transitions.

To compute the average transition time between two states, i.e., the average time it takes to reach state S_f when starting in state S_i, we employ the well-developed concept of mean first passage times (MFPTs) given as

$$\mathbb{E}[T_{i\to f}] = \begin{cases} 1 + \sum_j T_{i\to j}\,\mathbb{E}\big[T_{j\to f}\big] & i \neq f \\ 0 & i = f \end{cases}. \tag{1.54}$$

This equation can be rewritten [139] in a linear system of equations

$$\sum_{j \neq f} W_j^i \, \mathbb{E}\left[T_{j \to f}\right] = -1 \quad \text{for } i \neq f \tag{1.55}$$

when using transition rates. This linear system of equations is then solved for the vector $\left\{\mathbb{E}\left[T_{j \to f}\right] \;\middle|\; j \neq f\right\}$ containing the requested mean first passage time $\mathbb{E}[T_{i \to f}]$. To compute, more generally, the MFPT between two disjunct sets of states, say, A and B with $A \cap B = \emptyset$, eq. (1.54) is extended to

$$\mathbb{E}[T_{A \to B}] = \sum_{a \in A} \frac{\pi_a \, \mathbb{E}[T_{a \to B}]}{\sum_{k \in A} \pi_k}, \tag{1.56}$$

where MFPTs $\mathbb{E}[T_{a \to B}]$ are weighted by their relative steady-state probabilities $\pi_a / \sum_{k \in A} \pi_k$. An illustration of an exemplary graph $\mathcal{G}(V, E, W)$ is shown in Figure 1.2. The MFPT from the subset A (red shaded area) to the target set B (blue shaded area) corresponds to the average time it takes a process originated in A to first hit the target set B.

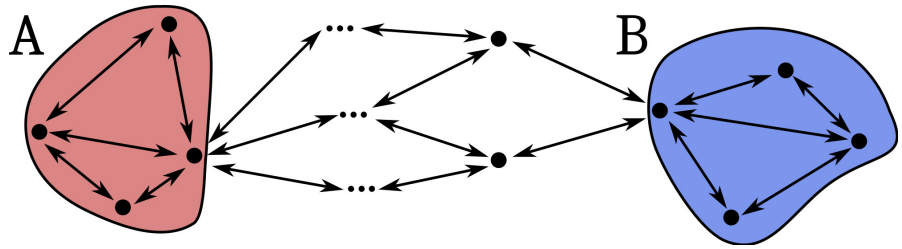

Figure 1.2 – **Transition network.** Illustration of a graph $\mathcal{G}(V, E, W)$ with vertices V indicated by filled circles and directed edges by double-headed arrows. Different edge weights (transition rates) are not accounted for. Nonoverlapping subgraphs are highlighted by red (A) and blue (B) shaded areas. Average transition times, e.g., between set A and B can be determined by computing mean first passage times, see eq. (1.56).

1.3 Coarse graining of equilibrium MSMs

For the construction of MSMs the configuration space is usually fine grained by many hundreds or even thousands of centroids (micro states), especially if no prior information about the free energy landscape is available. The resulting fine-grained transition matrix does, in principle, hold all information which are required to describe the long-term dynamics. However, it is a cumbersome task to interpret a dynamics that includes hundreds or even thousands of states. In the previous section, we showed how to compute transition times between a given initial and target set of states. The question remains how to determine these sets when not known beforehand. Our task is thus to coarse grain an estimated transition matrix into an effective one with as few states as possible, while still being able to describe the original long-term dynamics. In this section, we deal with equilibrium MSMs (obeying detailed balance), whereas the coarse graining of nonequilibrium MSMs is discussed in chapter 2.

The first step of coarse graining a MSM is to partition all N states of the transition network – represented by the graph $\mathcal{G}(V,E,T)^\star$ – into k clusters with $k \ll N$. Many different algorithms have been proposed detecting "closely related" vertices and grouping them into clusters, sometimes referred to as communities. These "closely related" vertices are often thought to share many edges or edge weights, although different conditions exist [49]. Besides the detailed working principle, graph partitioning algorithms can be categorized into two classes: algorithms which use the number of clusters k as input and algorithms which determine the number of clusters dynamically.

Here we discuss the **Robust Perron-Cluster Cluster Analysis** (PCCA+) [114] a spectral clustering algorithm which falls into the first category, i.e., the number of clusters k needs to be known beforehand.

Suppose the eigenvalue spectrum of the transitions matrix $\boldsymbol{T}(\tau)$ offers a gap between its kth and $(k+1)$th eigenvalue, i.e., $1 = \Lambda_0 > \Lambda_1 > ... > \Lambda_k \gg \Lambda_{k+1} > ... > \Lambda_{N-1} \approx 0$, then $\boldsymbol{T}$ can be expanded into k slow and $(N-k)$ fast decaying modes, analogous to eq. (1.28). The PCCA+ algorithms exploits the sign structure of the first k (right) eigenvectors to partition the graph $\mathcal{G}(V,E,T)$ into $k+1$ clusters or metastable sets by solving a linear system of equations. To this end, every vertex is associated with every coarse-grained set by an individual probability (fuzzy partitioning). The advantage of fuzzy partitioning is that some vertices cannot be unambiguously assigned to a specific coarse-grained set (as in crisp partitioning), in the sense that the average time it takes to leave the coarse-grained or metastable set is much longer than the average time of staying within.

If the distribution of all k sets is known, the coarse-grained transition rate matrix $\boldsymbol{W}_{\mathrm{cg}} \in \mathbb{R}^{k\times k}$ representing the dynamics between these metastable sets, is obtained by computing inverse mean first passage times,

$$(W_{\mathrm{cg}})_J^I = \mathbb{E}\left[T_{I\to J}\right]^{-1}. \tag{1.57}$$

1.4 Paradigm of rate estimation: one-dimensional particle in a double well potential

After setting the stage for constructing and analyzing equilibrium Markov state models, we give an illustrative example: an one-dimensional Brownian particle in a double well potential. To describe the particle dynamics, we employ the overdamped Langevin equation [see section 1.1.2] reading

$$\dot{x} = -\frac{\nabla U(x)}{\gamma_{\mathrm{fric}}} + \eta(t) \qquad \text{with} \qquad U(x) = \varepsilon\left(\frac{1}{2}x^4 - 4x^2\right). \tag{1.58}$$

*Here $\mathcal{G}(V,E,T)$ refers to a graph with edge weights given by the transition probabilities.

An exemplary trajectory, which illustrates stochastic transitions between both potential energy minima, is depicted in Figure 1.3(a).

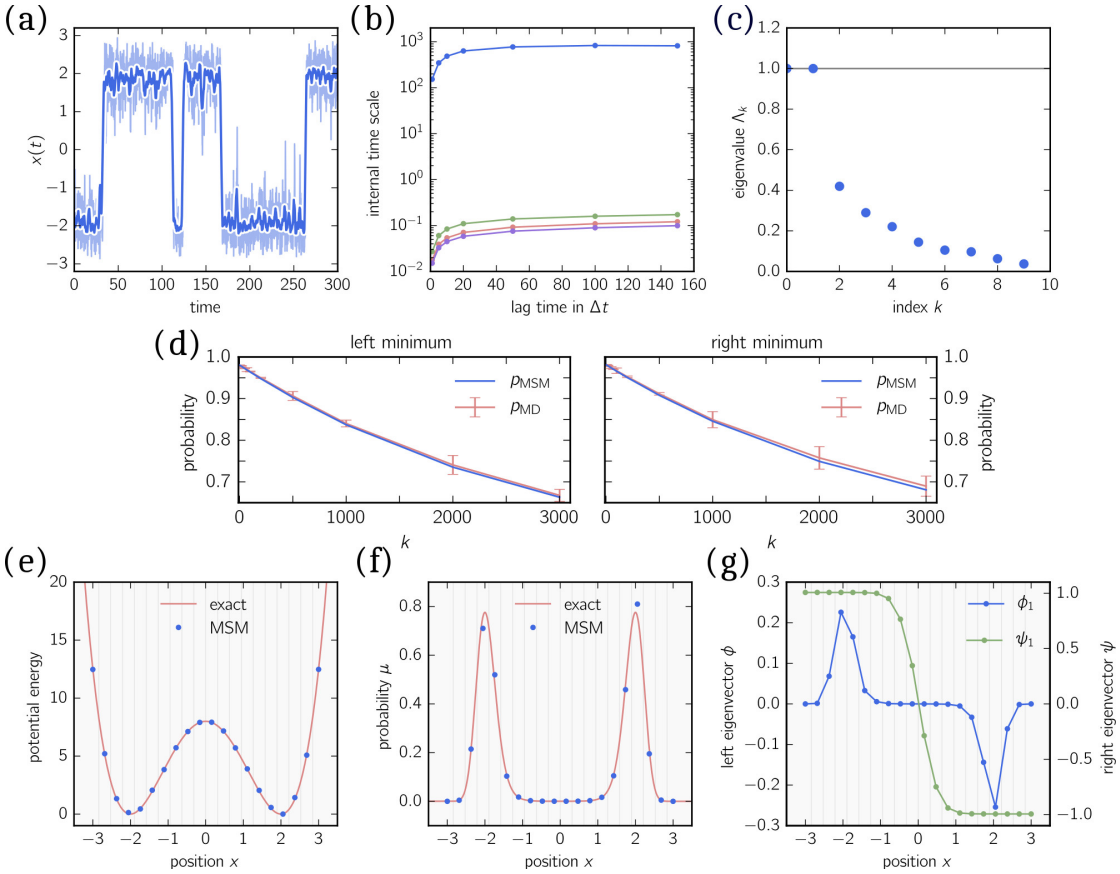

Figure 1.3 – **Double well potential.** (a) Exemplary time series illustrating stochastic transitions between both potential wells. The shaded time series depicts the full trajectory, while the solid line indicates its moving average ($\tau = 2$). Employed simulation parameters are $D = 1$ and $\epsilon = 0.5\gamma_{\text{fric}}$. (b) Internal time scales are shown for multiple lag times τ. (c) Eigenvalue spectrum of the transition matrix $\boldsymbol{T}(\tau = 100\,\Delta t)$. (d) Chapman-Kolmogorov test with initial probability distribution located in the left and right potential energy minimum. (e) True potential energy landscape (red) and MSM approximation (blue). (f) True probability distribution (red) and MSM approximation (blue). (g) Approximated first left and right eigenvector. (e-g) Gray areas illustrate the discretized position space, while x values of the blue points indicate centroid positions.

To sample the potential energy landscape, multiple trajectories are recorded accumulating a total of $T = 200$ (Brownian time units). The position space is discretized by 20 equally placed centroids, which are illustrated by blue points (centroid positions) and gray intervals (discrete states) in Figure 1.3(e). Imposing the detailed balance condition, the transition matrix $\boldsymbol{T}(\tau)$ is estimated for different lag times τ by employing the maximum likelihood estimator. The lag time analysis, see Figure 1.3(b), suggests that for lag times $\tau \geq 100\,\Delta t = 0.1$ the slowest implied time scale becomes lag time independent. As the second slowest implied time scale is already

more than three orders of magnitude faster, we expect the eigenvalue spectrum to show a gap between Λ_1 and Λ_2. Indeed, as shown in Figure 1.3(c), the second largest eigenvalue Λ_1 is almost one, while $\Lambda_2 \approx 0.4$.

In order to test for Markovianity of the slow dynamics we perform the Chapman-Kolmogorov test as introduced in section 1.2.3. We test for two different probability distributions $\boldsymbol{\omega}^A$ located in either the left or right minimum. As shown in Figure 1.3(d), the time evolutions of both $\boldsymbol{\omega}^A$'s that follow the MSM are in very good agreement with the probability distributions determined from the original data set.

In Figure 1.3(e) and (f), we show the analytical as well as reconstructed potential energy landscape and equilibrium probability distribution, respectively. To compare both probability distributions we normalize the curves within the shown interval, i.e., $x \in [-3,3]$, whereas the reconstructed potential energy follows from $U^{(i)}_{\text{pot}} = -\log \pi_i$. For both quantities the MSM solution in very good agreement with the true curve. Solely, the height of the approximated probability distribution exhibits small deviations from the analytical curve, which becomes negligible for the potential energy landscape.

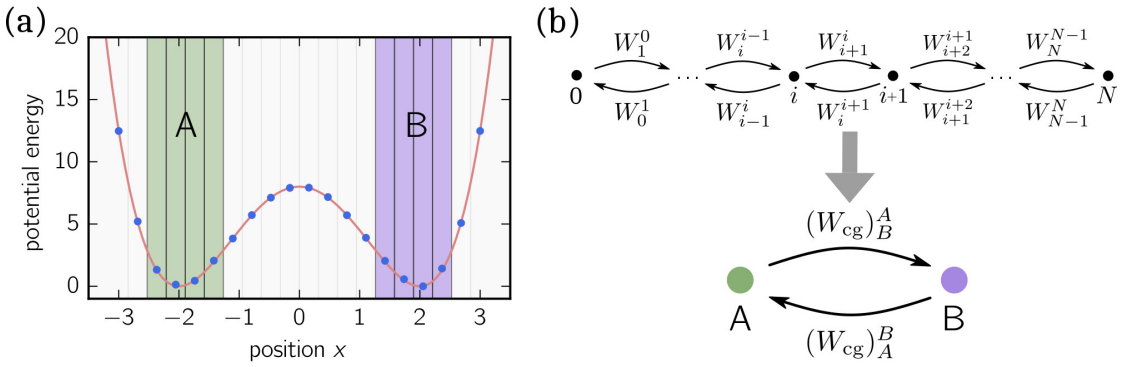

Figure 1.4 – **Coarse-grained MSM.** (a) Potential energy landscape as shown in Figure 1.3(e). Blue and green colored intervals indicate metastable sets, while gray intervals indicate transition states. (b) Transition network according to the full and coarse-grained Markov state model with edge weights representing transition rates. The states constituting the metastable sets shown in (a) are lumped together forming the coarse-grained states as shown (b). Coarse-grained transition rates are computed by inverse mean first passage times.

The large time scale separation as indicated by the lag time analysis [Figure 1.3b] suggests to coarse grain the graph $\mathcal{G}(V,E,T)$ into two metastable sets with the slowest implied time scale corresponding to the transition between both metastable sets. Figure 1.3(g) depicts the left and right eigenvector ϕ_1/ψ_1 with $\psi_1 = \phi_1\,\pi$, clearly illustrating that the sign structure of ψ_1 can be exploited to separate both potential minima, while the inflection point $\psi_1 \approx 0$ points to the top of the barrier of the double well potential. Applying the PCCA+ algorithm for $k = 1$, two metastable sets are identified, as shown in Figure 1.4(a), for which we identify transition states as states that cannot be assigned with a probability larger than 80% to one of the metastable sets.

Finally, to coarse grain the dynamics we compute inverse mean first passage times, see eq. (1.57), between states belonging to set A and B. In Figure 1.4(b), we illustrate the graph indicating the full and coarse-grained dynamics.

1.5 Stochastic thermodynamics of discrete systems

While we discussed the concept of Markov state modeling limited to systems in thermal equilibrium in the previous section, we consider here systems that are driven in a **nonequilibrium steady state** (NESS). For a NESS, in contrast to thermal equilibrium, the dynamics (or more precisely the associated phase space propagator or generator) breaks detailed balance. In this regard, we understand thermal equilibrium as "just" a special form of a steady state and a NESS as its logical extension. Before discussing the consequences for a system being in a NESS, we give an illustrative example pointing out the differences to thermal equilibrium.

The molecular motor protein kinesin is one of the best studied biochemical systems, responsible for cellular functions such as mitoses and meiosis as well as the transport of cellular cargo [62, 120]. The kinesin complex comprises multiple subunits which are schematically shown in Figure 1.5(a). The "feet" (motor domain) are the active binding sites of the protein and constitute the two heads of the protein which are linked via short, flexible molecular linkers to the body or stalk. The body is connected to the tail component which is responsible for attaching molecular cargo. Each head component contains two separate binding sites: one binding site for attaching the head to a microtubule, which is an intracellular filament also known as "cellular highway", and one binding site for binding andenosine triphosphate (ATP).

When ATP is hydrolyzed to adenosine diphosphate (ADP) and inorganic phosphate (P), chemical energy is released which triggers a conformational change of the microtubule-binding domain, propelling the head forward along the microtubule (mechanical transition). From a physical point of view, the chemical reaction ATP $\rightarrow$ ADP+P releases chemical energy which is converted into mechanical work as the kinesin protein moves along the microtubule.

To keep the hydrolysis rate constant and thus the kinesin protein walking, the chemical potential of ATP has to be kept constant, which requires the coupling to an ATP reservoir. Because the synthesis of ATP costs energy, we understand that to maintain the kinesin motor in a steady state, the motor protein or rather its environment must be constantly supplied with (chemical) energy. This energy influx is identified as the driving mechanism that separates thermal equilibrium (no energy influx) from a NESS.

In the study of Liepelt *et al.* [83], the authors propose a minimal **experimentally probed** Markov model elucidating the chemo-mechanical mechanism of the kinesin motor protein. The model, shown in Figure 1.5b, includes six different states that are characterized by both kinesin heads (active sites). Each head can be occupied by an ATP, ADP or no molecule, yielding six different combinations which are referred to as states of the model. For example, in state 1 an ADP molecule is bound to the blue head while the red head is unoccupied. If two

states are connected, indicated by arrows in Figure 1.5(b), transitions between both states are possible, implying nonzero transition rates W_j^i. For every connected pair of states forward and backward transition rates are either experimentally measured or deduced, e.g., $\{W_1^0, W_0^1\} > 0$. While the solid black arrows indicate chemical transitions, i.e., molecules are released, newly bound or chemically transformed, the dashed arrow $2 \leftrightarrow 5$ indicates the mechanical transition, that is, the kinesin molecule makes a forward or backward step, respectively.

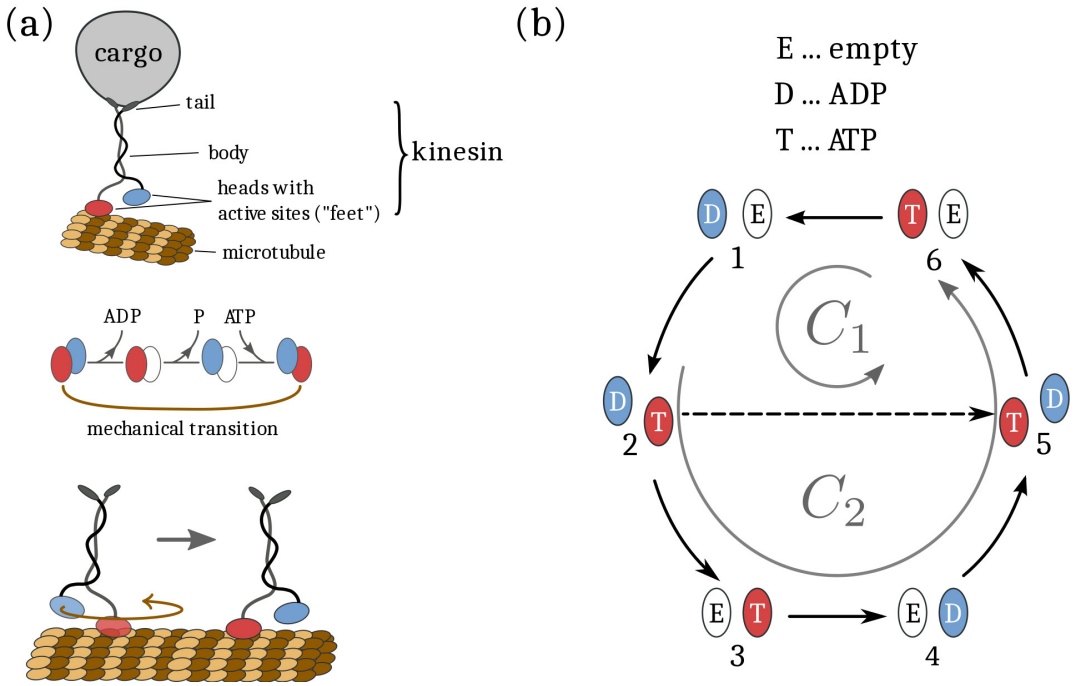

Figure 1.5 – **Kinesin network.** (a) Sketch of structure and function of the kinesin molecule. The kinesin molecule converts chemical energy in the form of ATP (andenosine triphosphate) into mechanical work, causing the kinesin molecule to "walk" along cellular microtubules. (b) Transition network of kinesin dynamics. While states (labeled by numbers) indicate different occupations of binding sites, black arrows indicate the flow direction of probability currents within the transition network. Moreover, the dashed arrow shows the mechanical transition (i.e., the kinesin molecule makes a step forward), and gray arrows illustrate the probability current cycles found in the network. The figure is adopted from ref. [6].

As stated above, in contrast to the example given in the previous section, the dynamics of the kinesin network breaks detailed balance as the system constantly dissipates energy. The deviation from detailed balance

$$W_j^i \pi_i - W_i^j \pi_j = \Phi_j^i - \Phi_i^j \equiv J_j^i \tag{1.59}$$

defines **probability currents** J_j^i which are the difference between their respective forward and backward **probability fluxes** Φ_j^i. The probability currents indicate the net transport of probability from states $i \to j$. Thus, if the probability current J_j^i is positive, a trajectory jumps more often from state $i \to j$ than from state $j \to i$. It is important to note that physical quantities

such as charge, particles, etc., can only be transported between states if the respective probability current is nonzero. More generally stated, in thermal equilibrium the thermodynamic average of every current variable X vanishes, i.e., $\langle X \rangle = 0$, because the associated probability distribution is symmetric, i.e., $\mathbb{P}(-X) = \mathbb{P}(X)$.

Having defined probability currents, the element-wise master equation can be rewritten, revealing a useful property known as **Kirchhoff's current law**

$$0 = \partial_t \pi_i = \sum_j W_i^j \pi_j = W_i^i \pi_i + \sum_{j \neq i} W_i^j \pi_j \overset{W_i^i = -\sum_{j \neq i} W_j^i}{=} \sum_{j \neq i} \left[W_i^j \pi_j - W_j^i \pi_i \right] = -\sum_{j \neq i} J_j^i \tag{1.60}$$

that states that the same amount of probability that flows into a state also flows out.

Following the direction of the positive probability currents, as shown in Figure 1.5(b) (indicated by arrow heads), reveals another interesting property of a NESS. Because probability is a conserved quantity, that is, no probability sinks or sources exist, probability must be transported in **cycles**. We define these probability current cycles as an ordered set of states (vertices), at the end of which the starting state is reached again and no other vertex occurs twice. Cycles that differ only in their cyclic permutation of states are considered identical. For instance, $\{1,2,3,1\} = \{2,3,1,2\} = \{3,1,2,3\}$ all denote the same cycle but $\{3,2,1,3\}$ is a different cycle.

The kinesin network exhibits two different cycles: the forward cycle $C_1 = \{1,2,5,6,1\}$, because after completion the kinesin protein has made a step forward, and the dissipative cycle $C_2 = \{1,2,3,4,5,6,1\}$, in which the kinesin does not move, although chemical energy is dissipated.

So far we have loosely argued that a NESS is maintained by constantly supplying the system with some form of energy. Following the ideas of irreversible thermodynamics [26], we introduce the concept of **generalized thermodynamic forces** which are often referred to as **affinities**. These generalized forces express the mechanisms responsible for driving a system out of equilibrium, while physical currents are the system's response. Since the same concept holds for macroscopic systems, we illustrate this idea by a macroscopic example.

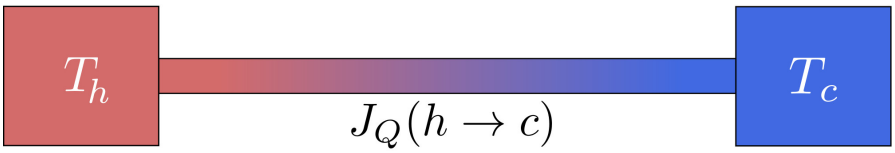

Figure 1.6 – **Schematic illustration of a hot and cold heat reservoir connected by a heat conducting wire.** The system constantly transports heat from the hot to the cold reservoir.

Suppose a hot and cold heat reservoir are connected by a heat conducting wire, as schematically depicted in Figure 1.6. In the steady state the wire exhibits a constant temperature

gradient that generates a constant heat current J_Q flowing from the hot to the cold reservoir. The affinity responsible for the heat current is given by the inverse temperature gradient $f = T_c^{-1} - T_h^{-1}$. Generally, currents can be expressed as nonlinear functions of their respective affinity. However, when affinities are small, i.e., the NESS is within the linear response regime, currents are proportional to their respective affinities with their proportionality factors known as transport coefficients [26].

As the second law of thermodynamics states that the amount of entropy can never decrease, the average entropy production rate in a NESS is identified [26] as

$$\dot{S} = \sum_k f_k J_k \geq 0. \tag{1.61}$$

Here f_k denotes the kth affinity (generalized force) and J_k its corresponding current. For a system in thermal equilibrium all affinities and currents vanish (detailed balance holds) and thus the total entropy of the system stays constant.

The generalization of these concepts to microscopic systems is given by the framework of stochastic thermodynamics [125]. The crucial difference to macroscopic systems is that for microscopic length scales fluctuations become critical, which are accounted for by the concept of fluctuation theorems [67, 34, 43] that are valid even far beyond the linear response regime. Typically, the framework of stochastic thermodynamics is applied to describe systems, e.g., colloidal systems, biochemical systems, etc., that are immersed in a heat bath referred to as **environment or medium**. The environment is implicitly described by Gaussian noise as introduced in the framework of Langevin dynamics [see section 1.1.2]. Although the theoretical concepts of stochastic thermodynamics are derived for space-continuous systems, they equally apply to space-discrete systems described by the master equation, and therefore extend the early ideas of Schnakenberg [121] and Hill [60].

Similar to the description of macroscopic thermodynamic systems, the first law of thermodynamics can be applied to microscopic transitions, i.e.,

$$(\Delta U)^i_j \equiv U(j) - U(i) = w^i_j - q^i_j, \tag{1.62}$$

where $U(i)$ is the potential energy of state i, w^i_j the performed work and q^i_j the dissipated heat for the transition $i \to j$. Next, one assumes that the work per transition is given by

$$w^i_j = f\, d^i_j, \tag{1.63}$$

where f denotes the generalized force [or macroscopic affinity, see eq. (1.61)] and d^i_j the generalized distance for the transition $i \to j$, which is consequently anti symmetric, i.e., $d^i_j = -d^j_i$ [129, 125]. The generalized distances correspond to transported (or produced) quantities on the microscopic level. For example, for a sheared fluid the generalized force corresponds to the stress, while the generalized distance d^i_j expresses the amount of strain "produced" for the transition $i \to j$.

When starting from the opposite direction, that is, all microscopic rates and probabilities are known, **microscopic affinities** are defined as

$$A_j^i \equiv k_B \log\left(\frac{\Phi_j^i}{\Phi_i^j}\right) = k_B \left[\log\left(W_j^i \pi_i\right) - \log\left(W_i^j \pi_j\right)\right], \tag{1.64}$$

which have the physical meaning of the total entropy produced for the transition $i \to j$, i.e., $A_j^i \equiv (\Delta S_{\text{tot}})_j^i$. Employing the **entropy balance equation** [126, 45], the total produced entropy per transition is split into two terms:

$$(\Delta S_{\text{tot}})_j^i = (\Delta S_{\text{med}})_j^i + (\Delta S_{\text{sys}})_j^i, \tag{1.65}$$

where $(\Delta S_{\text{med}})_j^i \equiv \sigma_j^i$ expresses the entropy produced in the medium (environment) and $(\Delta S_{\text{sys}})_j^i$ the entropy produced in the system.

Connecting microscopic affinities [eq. (1.64)] with the entropy balance equation [eq. (1.65)]

$$A_j^i = k_B \left[\log\left(W_j^i \pi_i\right) - \log\left(W_i^j \pi_j\right)\right] = \sigma_j^i + (\Delta S_{\text{sys}})_j^i, \tag{1.66}$$

and plugging in the definition of the system's entropy

$$(\Delta S_{\text{sys}})_j^i = S_{\text{sys}}(j) - S_{\text{sys}}(i) = -k_B \log(\pi_j) + k_B \log(\pi_j) = k_B \log\left(\frac{\pi_i}{\pi_j}\right), \tag{1.67}$$

we identify the entropy produced in the medium per transition as

$$\sigma_j^i = k_B \log\left(\frac{W_j^i}{W_i^j}\right). \tag{1.68}$$

We can further conclude that for thermal equilibrium ($A_j^i \equiv 0$) the medium has the function of an "entropy accountant". If, for example, the transition $i \to j$ produces entropy in the system, the same amount is depleted from the medium and vice versa. Moreover, from eq. (1.64) and eq. (1.68) it can be seen that microscopic affinities as well as medium entropies are anti symmetric, e.g., $A_j^i = -A_i^j$ and $\sigma_j^i = -\sigma_i^j$, and interestingly do not depend on the system's time scales.

Unfortunately, up to today a comprehensive theory that fully connects microscopic and macroscopic affinities is still missing. However, the concept of local detailed balance establishes a connection between macroscopic affinities and microscopic rates.

1.5.1 From macroscopic affinities to microscopic rates

In thermal equilibrium, one can exploit detailed balance and the Boltzmann distribution to gain an expression that connects the ratio of microscopic forward and backward rates, i.e.,

$$\frac{W_j^i}{W_i^j} = \frac{\pi_j^{\mathrm{eq}}}{\pi_i^{\mathrm{eq}}} = \exp\Big(-\beta\,(\Delta U)_j^i\Big) \quad \text{(thermal equilibrium)}. \tag{1.69}$$

For a NESS, this expression can be extended to

$$\frac{W_j^i}{W_i^j} = \frac{\pi_j^{\mathrm{eq}}}{\pi_i^{\mathrm{eq}}}\,\exp\Big(\beta\, f d_j^i\Big) \quad \text{(NESS)}, \tag{1.70}$$

which is known as the **local detailed balance** condition [see for instance refs. [74, 125, 44]]. Here π_i^{eq} represents the thermal equilibrium probability of state i [see eq. (1.69)], that is, the probability to occupy state i if all microscopic affinities would vanish. Moreover, from the local detailed balance condition, the first law of thermodynamics [see eq. (1.62)] and eq. (1.63) we identify the dissipated heat with the entropy produced in the environment at temperature T,

$$\sigma_j^i = k_B\,\log\left(\frac{W_j^i}{W_i^j}\right) = \beta\, f d_j^i - \beta\,(\Delta U)_j^i = \beta\, q_j^i. \tag{1.71}$$

1.5.2 Entropy production rate for discrete systems

To connect the microscopic affinities with the macroscopic entropy production rate, cf. eq. (1.61), we sum the microscopic affinities over all transitions weighted with their respective probability current, i.e.,

$$\langle \dot{S}_{\mathrm{tot}}\rangle = \langle A\rangle = \frac{1}{2}\sum_{ij} J_j^i\, A_j^i = \underbrace{\frac{1}{2} k_B \sum_{ij} J_j^i\,\log\left(\frac{\pi_i}{\pi_j}\right)}_{\langle \dot{S}_{\mathrm{sys}}\rangle} + \langle\sigma\rangle. \tag{1.72}$$

Together with eq. (1.65) the first term one the right hand side of eq. (1.72) is identified as the entropy production rate of the system and the second term as the medium entropy production rate. One should note that, in agreement with the second law of thermodynamics, only $\langle \dot{S}_{\mathrm{tot}}\rangle$ is larger than zero.

A second but intriguing approach to derive the entropy production rate of the medium is to start with the definition of the Gibbs entropy that describes the entropy of the system [126]

$$S_{\mathrm{Gibbs}}(t) = \langle S_{\mathrm{sys}}(t)\rangle = -k_B \sum_i p_i(t)\log\big(p_i(t)\big). \tag{1.73}$$

Its time derivative reads

$$\begin{aligned}
\frac{\mathrm{d}S_{\mathrm{Gibbs}}(t)}{\mathrm{d}t} &= -k_B \sum_i \Big(\partial_t p_i\Big)\log\big(p_i\big) + k_B \sum_i \frac{\partial_t p_i}{p_i} \\
&\overset{\text{eq. (1.35)}}{=} -k_B \sum_i \left[\sum_{j\neq i}\Big(W_i^j p_j - W_j^i p_i\Big)\right]\log\big(p_i\big) + k_B \sum_i \frac{\partial_t p_i}{p_i} \\
&= k_B \sum_{ij} W_j^i p_i \log\left(\frac{p_i}{p_j}\right) + k_B \sum_i \frac{\partial_t p_i}{p_i} \\
&\overset{p_i(t)=\pi_i}{=} k_B \sum_{ij} W_j^i \pi_i \log\left(\frac{\pi_i}{\pi_j}\right) = \frac{1}{2} k_B \sum_{ij} J_j^i \log\left(\frac{\pi_i}{\pi_j}\right) = \langle \dot{S}_{\mathrm{sys}}\rangle .
\end{aligned} \tag{1.74}$$

When rewriting the entropy balance equation [eq. (1.65)] for the mean entropy production rates and employing the expressions for $\langle \dot{S}_{\mathrm{sys}}\rangle$ and $\langle \dot{S}_{\mathrm{tot}}\rangle$, we identify the mean entropy production rate of the medium as

$$\langle \sigma\rangle = \langle \dot{S}_{\mathrm{tot}}\rangle - \langle \dot{S}_{\mathrm{sys}}\rangle = \langle A\rangle - \frac{1}{2} k_B \sum_{ij} J_j^i \log\left(\frac{\pi_i}{\pi_j}\right) = \frac{1}{2} k_B \sum_{ij} J_j^i \log\left(\frac{W_j^i}{W_i^j}\right). \tag{1.75}$$

If the system is in a NESS, $\langle \dot{S}_{\mathrm{sys}}\rangle$ contributes to the total entropy production rate only in form of a boundary term which vanishes for $t \to \infty$ and therefore $\langle \dot{S}_{\mathrm{tot}}\rangle = \langle A\rangle = \langle \sigma\rangle$.

2 Coarse graining of nonequilibrium Markov state models

Having established the connection with stochastic thermodynamics, the goal of this chapter is to derive a dynamically consistent coarse graining strategy allowing the construction and analysis of nonequilibrium Markov state models (NE-MSMs) from molecular dynamics simulations. To this end, we illustrate all steps by an easy but intuitive example: a particle trapped in a double-well potential, driven by a nonconservative force.

2.1 Cycle representation

2.1.1 Current- and state-like observables

Generally, we distinguish two types of observables: First, state-like observables, i.e., observables evaluated for the states (vertices) of the system and second, current-like observables, i.e., observables evaluated along the transitions (edges) that connect the states of the system. As both types of observables can have large fluctuations, we define their averages according to

$$\langle O\rangle = \sum_i \pi_i O_i \quad \text{(state-like)} \quad \text{and} \quad \langle O\rangle = \sum_{ij} \Phi^i_j O^i_j = \frac{1}{2}\sum_{ij} J^i_j O^i_j \quad \text{(current-like)}, \tag{2.1}$$

where state-like averages represent regular ensemble averages, whereas current-like averages express the average change of the observable per unit time. Note that if detailed balance holds all current-like averages become zero. In the previous section, we introduced the concept of cycles as an ordered set of states. We further distinguish between **trivial cycles**, i.e., cycles that contain two different states, e.g., $C = \{1,2,1\}$, and **nontrivial cycles** containing more than two different states.

For convenience, we define the indicator

$$\chi^i_\alpha = \begin{cases} 1 & \text{if vertex } i \text{ is in cycle } \alpha \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

and the passage function

$$\chi^i_{j,\alpha} = \begin{cases} 1 & \text{if directed edge } i \to j \text{ is in cycle } \alpha \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

depending on whether a vertex i or directed edge $i \to j$ is part of a cycle α.

To extend the definition of state- and current-like observables for cycles, we define

(i) State-like cycle observables that are summed over the states forming a cycle, i.e.,

$$O_\alpha = \sum_i O_i \chi^i_\alpha. \tag{2.4}$$

(ii) Current-like cycle observables that are summed along the edges corresponding to each cycle, i.e.,

$$O_\alpha = \sum_{ij} O^i_j \chi^i_{j,\alpha}. \tag{2.5}$$

To distinguish cycle observables from "regular" observables all cycle indices are denoted by Greek letters.

2.1.2 Cycle affinities

An important example for a current-like cycle observable is the **cycle affinity**

$$A_\alpha = \sum_{ij} A^i_j \chi^i_j, \tag{2.6}$$

which represents the total entropy produced when traversing the full cycle α. Cycle affinities have two crucial properties:

(i) The cycle affinity of trivial cycles is always zero since affinities are anti-symmetric, $A^i_j = -A^j_i$, and thus $A_{\{i,j,i\}} = A^i_j + A^j_i = A^i_j - A^i_j = 0$.

(ii) All cycle affinities are independent of the steady-state probability distribution and thus can be solely expressed in terms of σ^i_j's.

To prove (ii),

$$A_{\{i,j,\ldots,n,i\}} = A^i_j + \cdots + A^n_i = \log\left(\frac{W^i_j \pi_i \ldots W^n_i \pi_n}{W^j_i \pi_j \ldots W^i_n \pi_i}\right) = \log\left(\frac{W^i_j \ldots W^n_i}{W^j_i \ldots W^i_n}\right) = \sigma^i_j + \cdots + \sigma^n_i. \tag{2.7}$$

Since the π_i's are state functions, going along a cycle they cancel pairwise. A remarkable property of a NESS is therefore that the average entropy production can be calculated either

by using the A^i_j's or σ^i_j's.

2.1.3 Cycle averages

At this stage we have introduced cycles as an indispensable concept for transport properties. In particular, the transportation of probability is of utmost importance if one wants to understand the dynamics of a NESS. Due to the close connection between cycles and the transportation of probability, it is reasonable to ask if the averages of current-like observables can be computed by employing cycle observables. To verify this, we introduce the **cycle-flux decomposition**, a graph theoretical tool [73, 7] that expresses the probability flux matrix $\mathbf{\Phi}$ through a linear combination of cycles, i.e.,

$$\Phi^i_j = \sum_\alpha \varphi_\alpha \chi^i_{j,\alpha}, \tag{2.8}$$

where the sum is over all cycles α that can be found in graph $\mathscr{G}(V,E,\Phi)$. The nonzero coefficient φ_α is referred to as **cycle weight** and indicates the probability flux flowing through cycle α. Mathematically, the cycle-flux decomposition is proven to always exist for every connected directed graph [73] that satisfies Kirchoff's current law, see eq. (1.60). A numerical algorithm efficiently computing the cycle-flux decomposition is given in the next section.

Using the definition of current-like cycle observables given in eq. (2.5) and inserting the cycle decomposition [eq. (2.8)], current-like averages

$$\langle O\rangle = \sum_{ij} \Phi^i_j O^i_j = \sum_\alpha \sum_{ij} O^i_j \chi^i_{j,\alpha} = \sum_\alpha \varphi_\alpha O_\alpha \tag{2.9}$$

can be expressed as current-like cycle averages. This expression is well defined even if the values of the cycle weights φ_α are not unique.

Of particular importance is that the total entropy production rate, see eq. (1.72), can be expressed as the cycle average of affinities

$$\langle A\rangle = \sum_\alpha \varphi_\alpha A_\alpha = \sum_{ij} \Phi^i_j A^i_j = \langle \dot{S}_{\mathrm{tot}}\rangle. \tag{2.10}$$

Finally, we define the mean entropy production rate of a single cycle $s_\alpha \equiv \varphi_\alpha A_\alpha$, such that $\langle \dot{S}_{\mathrm{tot}}\rangle = \sum_\alpha s_\alpha$.

2.2 Cycle-flux decomposition

In the literature several types of algorithms have been proposed to accomplish the decomposition given in eq. (2.8). One example is the "method of derived chain" introduced in [68], which is stochastic in nature and has the advantage that the cycle weights φ_α are unique and

that they correspond to the mean number of passages through the corresponding cycle α. However, negative cycle affinities ($A_\alpha < 0$) may occur, which greatly complicates the coarse graining of NE-MSMs (as will become clear shortly). Moreover, the number of cycles used in the decomposition can be orders of magnitude larger than for the other approaches discussed below. Another important type of cycle decomposition, first mentioned by Schnakenberg [121], considers fundamental cycles that span a basis of the cycle space. Although the number of contributing cycles is as small as for the described algorithm below, the fundamental cycles are not unique and can have negative cycle entropies.

Here we employ a variant of the "cycle-flux" decomposition, which is described and proven by Kalpazidou [73]. In contrast to the aforementioned algorithms, the cycle-flux decomposition only returns cycles exhibiting nonnegative cycle affinities ($A_\alpha \geq 0$). Moreover, the cycle-flux decomposition is a deterministic algorithm that has a polynominal complexity in the number of vertices $|V|$, making it computationally affordable even for large graphs.

However, the decomposition (and thus the cycle weights φ_α) is not unique but rather depends on the initial cycle sequencing. Therefore, already a minor variation in the sequencing can lead to different cycle weights. In particular, some cycle weights become zero, while others become nonzero. In ref. [7] this arbitrariness is illustrated by applying the cycle-flux decomposition to the totally asymmetric simple exclusion process model.

2.2.1 Algorithm

To decompose the edge fluxes Φ^i_j into cycles we start by splitting the flux matrix into two parts, i.e.,

$$\mathbf{\Phi} = \mathbf{\Phi}^{\mathrm{db}} + \tilde{\boldsymbol{J}}, \tag{2.11}$$

where $\mathbf{\Phi}^{\mathrm{db}}$ is a symmetric (detailed balance) matrix and $\tilde{\boldsymbol{J}}$ a nonnegative (current) matrix. The current elements $\tilde{J}^i_j$ are related to the microscopical currents, as introduced in eq. (1.59), by

$$\tilde{J}^i_j = \begin{cases} J^i_j & \text{for } J^i_j \geq 0 \\ 0 & \text{for } J^i_j < 0. \end{cases} \tag{2.12}$$

The nonnegative current matrix is obtained as $\tilde{\boldsymbol{J}} = \mathbf{\Phi} - \mathbf{\Phi}^{\mathrm{T}}$ with all negative elements set to zero, while the symmetric matrix follows from $\mathbf{\Phi}^{\mathrm{db}} = \mathbf{\Phi} - \tilde{\boldsymbol{J}}$.

At this point, we identify all trivial cycles, i.e., cycles with only two different states ($i \to j \to i$), by the nonzero elements of $\mathbf{\Phi}^{\mathrm{db}}$. As cycle weights we then assign their corresponding entries $\varphi_{i\to j\to i} \equiv \Phi^{i\,\mathrm{db}}_j = \Phi^{j\,\mathrm{db}}_i$.

To decompose the current matrix $\tilde{\boldsymbol{J}}$, we introduce an algorithm that can be split into two parts. First, we identify nontrivial cycles – cycles with more than two different states – and secondly

we determine their cycle weights. Theoretically, the number of possible cycles N_{cycles} in a graph grows exponentially with the number of vertices. For instance, the number of possible cycles in an undirected graph is bounded by $B \leq N_{\text{cycles}} \leq 2^B$, where $B \equiv |E| - |V| + 1$ is known as the **Betti number** [142]. For large $|V|$, B has the leading order of $|V|^2$, causing N_{cycles} to be bounded from above by $2^{|V|^2}$.

However, if both algorithmic steps – searching for a nontrivial cycle and determine its cycle weight – are combined by running them alternately, the decomposition becomes computationally affordable even for a large number of states.

To detect a nontrivial cycle, we propose the following steps:

(1) Find the position of the largest element of $\tilde{\boldsymbol{J}}$, i.e., $\arg\max\left(\tilde{J}^i_j\right)$.

(2) Identify the shortest path (smallest number of transitions) from state j leading back to state i by only following the nonzero transitions of $\tilde{\boldsymbol{J}}$. This step can be efficiently realized by applying a breadth-first search [95].

(3) Return the nontrivial cycle, i.e., $C_\alpha = \left\{i \to j \to \text{found path}\right\}$.

The breadth-first search algorithm is a standard tool for exploring tree or graph data structures. When starting from an initial vertex it stores the directly neighboring vertices, i.e., vertices that are connected by a directed edge pointing toward them. From these neighboring vertices then all next neighboring vertices are explored and so forth. If a vertex has been already explored, it is ignored in the next step. The breadth-first search algorithm stops when all vertices are explored, returning a so-called tree structure, that is, all $|V|$ vertices are connected by $|V| - 1$ edges. Considering the cycle detection in step (2) the algorithm stops when the vertex i is explored for the first time.

To determine the corresponding cycle weight φ_α, we take all flux values along cycle α and determine their smallest value, which becomes the cycle weight φ_α, i.e.,

$$\varphi_\alpha \equiv \min\left\{\tilde{J}^i_j \chi^i_{j,\alpha}\right\}. \tag{2.13}$$

Summing up both steps, the final cycle-flux algorithm reads

(1) Determine all trivial cycles and their cycle weights

(2) Find a nontrivial cycle in $\tilde{\boldsymbol{J}}$

(3) Compute its cycle weight φ_α

(4) Update all $\tilde{J}^i_j$ by subtracting φ_α along α, $\tilde{J}^i_j = \tilde{J}^i_j - \varphi_\alpha \chi^i_{j,\alpha}$ and continue with the second step

(5) The algorithm stops when the residuum $\left|\left|\tilde{\boldsymbol{J}}_{\max}\right|\right|$ has become smaller than a threshold.

2.2.2 Number of cycles

General considerations [73] show that the maximum number of needed cycles is bounded by $N_{\text{cycles}} = |E| - |V| + 1$ with $|E|$ being the number of nonzero elements of $\mathbf{\Phi}$ and $|V|$ its rank. For the number of nontrivial cycles we follow the reasoning of ref. [7]. First, one considers all trivial cycles appearing in a reversible graph, which number is identical to $|E|/2$. This becomes clear as each trivial cycle includes always the forward and backward edge connecting two states. Next, one looks at the maximum number of contributing nontrivial cycles. As described below, the algorithm first subtracts all trivial cycles, leaving an irreversible graph $\mathcal{G}(V, E, \tilde{J})$ with $|E|/2$ edges.

At this stage, one considers the worst case where each nontrivial cycle that has a nonzero cycle weight removes the current along one edge only. Because in each iteration the updated currents $\tilde{J}^i_j$ always fulfill Kirchhoff's current law, each vertex must be connected by at least two edges. The probability current going into a state has to flow out. After the algorithm has assigned $|E|/2 - |V|$ nonzero weights, the remaining nonzero currents include $|V|$ states with $|V|$ edges. Again, following Kirchhoff's current law, each vertex must be connected by at least two edges and therefore all $|V|$ states have to be connected by one cycle, that in the last iteration removes the remaining current. Hence, the number of contributing nontrivial cycles is given by $|E|/2 - |V| + 1$. Adding the $|E|/2$ identified trivial cycles gives an upper bound for the number of contributing cycles which is equal to the first Betti number $B = |E| - |V| + 1$.

2.2.3 Nonzero cycle affinities

To show that all by the proposed algorithm determined cycles indeed exhibit nonnegative cycle affinities, we consider only the nontrivial cycles, because cycle affinities of trivial cycles are zero, cf. section 2.1.2.

Since all edge affinities A^i_j of the original "current" graph $\mathcal{G}(V, E, \tilde{J})$ are positive and directed edges cannot be reversed throughout the iterations of the cycle-flux algorithm, all determined cycles follow the original edge directions. Therefore, all A_α are strictly positive for all nontrivial cycles that are found and zero for all trivial cycles. Note that this implies $s_\alpha > 0$ (cycle entropy production rate) for the nontrivial and $s_\alpha = 0$ for the trivial cycles, respectively.

2.2.4 Example

Finally, we give an intuitive example illustrating the cycle-flux decomposition as described in section 2.2.1.

In Figure 2.1(a), a simple, reversible and connected graph $\mathcal{G}(V, E, \Phi)$ is shown with vertices $V = \{A, B, C, D\}$. The arrows represent the edges, while numbers denote probability fluxes flowing in and out of each state. For example, the probability flux $\Phi^A_B = 4$ and the reverse flux $\Phi^B_A = 1$. It is easy to check that Kirchhoff's current law is valid, e.g., the flux into state A equals

the flux flowing out, $\Phi_A^{\text{in}} = 5+1+1 = 7 = 1+4+2 = \Phi_A^{\text{out}}$.

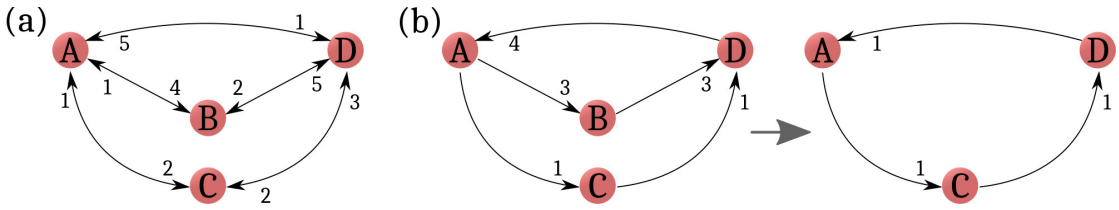

Figure 2.1 – **Example illustrating the cycle-flux decomposition.** The numbers next to the arrows are the numerical values for the probability fluxes Φ_j^i (a) and currents $\tilde{J}_j^i$ (b) along edges: (a) Initial graph $\mathcal{G}(V,E,\Phi)$ indicating probability fluxes, (b) current graph $\mathcal{G}(V,E,\tilde{J})$ before (left panel) and after (right panel) the first iteration of the cycle-flux decomposition.

According to the Betti number, the maximal number of cycles needed for the cycle-flux decomposition is seven ($B = 10-4+1 = 7$). In the first step of the decomposition all trivial cycles (the detailed balance part of the graph) are subtracted. All determined trivial cycles and their weights φ_α are listed in Table 2.1. The only possible nontrivial cycles that can exhibit nonzero cycle weights are $\{A,B,D,A\}$ and $\{A,C,D,A\}$, which are successively subtracted, cf. Figure 2.1(b).

Overall, seven cycles are required to complete the decomposition but only the two nontrivial cycles contribute to the mean entropy production rate, which is, according to eq. (2.10), given by $\langle \dot{S}_{\text{tot}} \rangle = 1 \cdot \log\ 15 + 3 \cdot \log\ 50 \approx 6.62$.

Table 2.1 – Cycles and cycle weights for the graph shown in Figure 2.1

Figure 2.1	cycles	weights φ_α	cycle affinity A_α
(a)	$\{A,B,A\}$	1	0
	$\{A,C,A\}$	1	0
	$\{A,D,A\}$	1	0
	$\{B,D,B\}$	2	0
	$\{C,D,C\}$	2	0
(b)	$\{A,B,D,A\}$	3	log 50
	$\{A,C,D,A\}$	1	log 15

2.3 Driven particle in a double well potential

Before we continue our task to coarse grain a nonequilibrium Markov state model, we introduce a two-dimensional model system: a particle trapped in a double well potential, driven out of equilibrium by a nonconservative force. Although the coarse graining algorithm that we will derive in this chapter is applicable to many-body systems, working in two dimensions allows to directly visualize the configuration space as well as the cycles within.

The system's dynamics is described by the overdamped Langevin equation [see section 1.1.2]

with the equation of motion for the position $\mathbf{x} = (x, y)^{\mathrm{T}}$ reading

$$\dot{\mathbf{x}} = -\frac{\nabla U}{\gamma_{\mathrm{fric}}} + \mathbf{f}_{\mathrm{nc}} + \boldsymbol{\eta}(t), \tag{2.14}$$

where $\boldsymbol{\eta}(t)$ is a random force with correlations $\langle \boldsymbol{\eta}_x(t)\boldsymbol{\eta}_y(t')\rangle = \sqrt{2D}\,\delta(t-t')\,\delta_{xy}$ and

$$U(x, y) = \varepsilon\left(\frac{x^4}{4} + \frac{1}{2}\left(y^2 - x^2 + x^2y^2\right)\right) \tag{2.15}$$

is the potential energy. As nonconservative force we choose $\mathbf{f}_{\mathrm{nc}}(\mathbf{x}) = \xi\,(-y, x)^{\mathrm{T}}$, where ξ denotes the driving strength. For the parameter values, we select $\varepsilon/\gamma_{\mathrm{fric}} = 20$, $\xi = 8$ and $D = 1$.

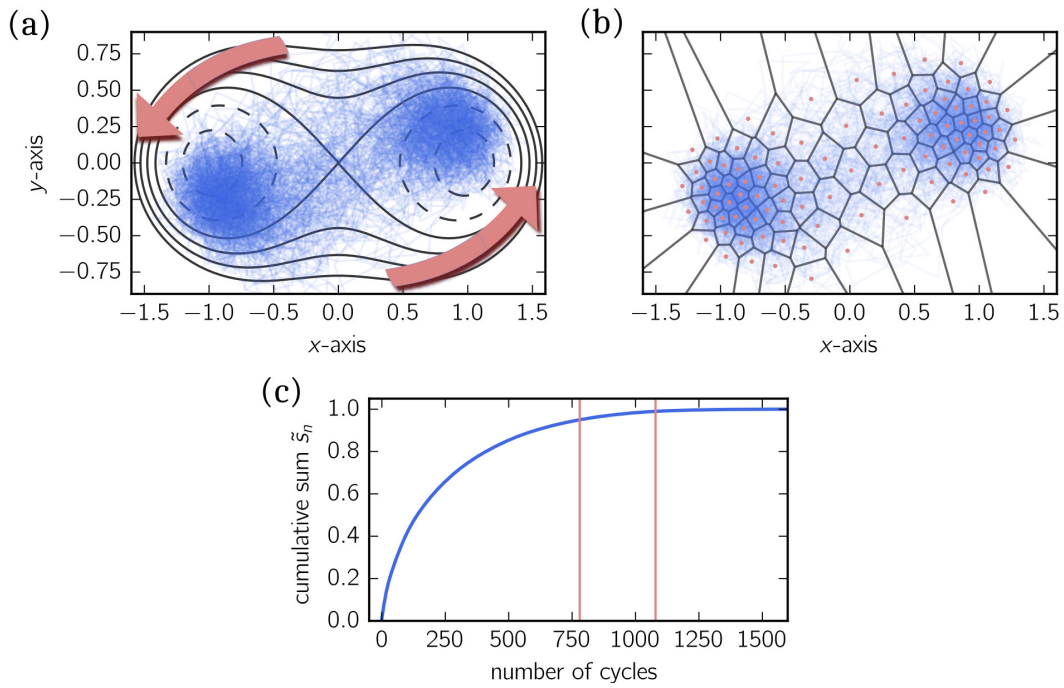

Figure 2.2 – **Driven particle in a double well potential.** (a) Contour lines of the symmetric two-dimensional potential energy as defined in eq. (2.15) and exemplary trajectory in the presence of the nonconservative force. Red arrows indicate the direction of the nonconservative force. The system is still invariant under inversion with the loci of highest probability being shifted from the minima of the potential. (b) Voronoi partition of the configuration space (cells are indicated by black lines) following $k = 100$ centroids determined by the k-means algorithms. Centroid positions are highlighted by red points. (c) Cumulated sum of cycle entropy production rates. Red lines indicate the number of cycles needed to recover 95% and 99% of the total entropy production rate, respectively.

Figure 2.2(a) shows the contour lines of the potential energy and an exemplary trajectory. Due to $\mathbf{f}_{\mathrm{nc}}$, the particle trajectory does not obey the symmetry of the conservative potential anymore.

To construct a NE-MSM from trajectories following the dynamics given by eq. (2.14), the configuration space is discretized by $k = 100$ centroids employing the k-means algorithm. The associated Voronoi decomposition is shown in Figure 2.2(b), where states (cells) are indicated by black lines and centroid position $\mathbf{R}_k$ by red dots. Following a lag time analysis, we estimate the transition matrix $\boldsymbol{T}(\tau = 0.07)$ by making use of the maximum likelihood solution, as introduced in eq. (1.39). The transition matrix is converted to the final rate matrix $\boldsymbol{W}$ as proposed in section 1.2.4.

At this point, we determine all trivial and nontrivial cycles by employing the cycle-flux decomposition [eq. (2.8)]. The graph $\mathcal{G}(V, E, \Phi)$ that illustrates the probability flux network consists of $|V| = 100$ vertices and $|E| = 3744$ edges. According to the Betti number $B = 3645$, we expect to find 1872 trivial and at most 1773 nontrivial cycles. In fact, the flux-cycle algorithm determines 1752 nontrivial cycles before it terminates.

We expect the number of nontrivial cycles to be even larger for NE-MSMs constructed from a configuration space that is discretized by many hundreds or even thousands of states. This is in stark contrast to the semiempirical NE-MSM for the kinesin protein that we introduced earlier. It is therefore undesirable to examine each cycle individually. In ref. [108] the authors conclude, after investigating artificial Markov models, which are not constructed from a physical dynamics, that "important" cycles can be recognized by their entropy production rate. Here, however, this conclusion fails as we do not find any dominant cycles that are responsible for a large fraction of the mean entropy production rate. To show this, we define the normalized cumulative sum

$$\tilde{s}_n \equiv \frac{\sum_{\alpha \leq n} s_\alpha}{\sum_\alpha s_\alpha} \tag{2.16}$$

of the first n nontrivial cycles for sorted cycle entropies $s_0 \geq s_1 \geq \ldots$. In Figure 2.2(c), $\tilde{s}_n$ is shown for the determined nontrivial cycles. The cumulative sum of the first 100 cycles does not even recover 30% of $\langle \dot{S}_{\mathrm{tot}} \rangle$. To reduce the number of nontrivial cycles, we consider only the 780 nontrivial cycles contributing to 95% of $\langle \dot{S}_{\mathrm{tot}} \rangle$.

2.4 Cycle space and communities

In the following our approach to coarse grain NE-MSMs is based on the idea that many cycles are similar in their length, traversed states (region in configuration space), and cycle affinities. Otherwise stated, the nontrivial cycles detected by the cycle-flux decomposition are not randomly positioned in configuration space but rather exhibit an ordered structure. We will propose how to quantify a "similarity" between cycles and how to group these cycles together, forming **cycle communities**.

To this end, we define **cycle centers** or rather their ith component as

$$c_\alpha^{(i)} \equiv \left(\frac{1}{|\alpha|} \sum_k \mathbf{R}_k^{(i)} \chi_\alpha^k \right), \tag{2.17}$$

where $|\alpha| = \sum_k \chi_\alpha^k$ denotes the length (number of states) of cycle α and $\mathbf{R}_k^{(i)}$ the ith component of the kth centroid. Furthermore, we define **cycle diameters** as

$$d_\alpha^{(i)} \equiv \max_{k,l} \left\{ \chi_\alpha^k \chi_\alpha^l \left| \mathbf{R}_k^{(i)} - \mathbf{R}_l^{(i)} \right| \right\}. \tag{2.18}$$

The cycle diameters indicate the largest distances spanned by the cycle vertices, as illustrated in Figure 2.3(a).

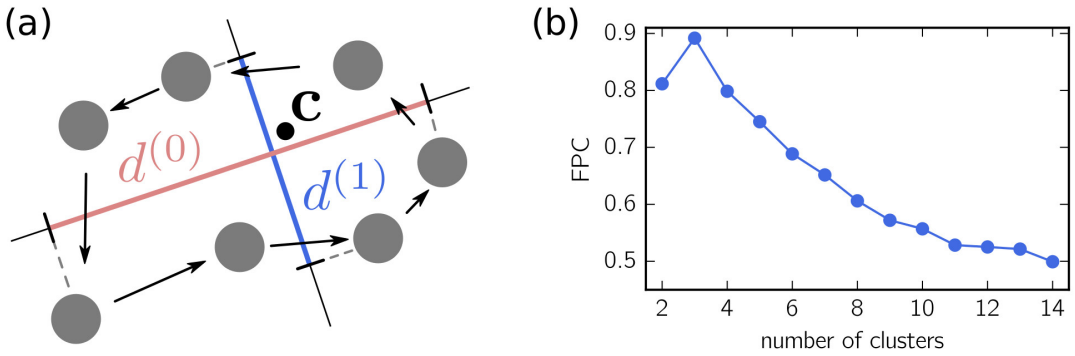

Figure 2.3 – **Cycle clustering.** (a) Schematic illustration of a nontrivial cycle, where gray circles denote vertices and arrows directed edges. Cycle diameters $d^{(0)}, d^{(1)}$ are obtained by projection of vertex (centroid) coordinates. The black point denotes the cycle center $\mathbf{c} = (c^{(0)}, c^{(1)})$. (b) The fuzzy partition coefficient (FPC) is shown for multiple cycle communities. The best result is obtained for three communities.

In general, if the configuration space $\Gamma \in \mathbb{R}^n$, then $\mathbf{c}_\alpha, \mathbf{d}_\alpha \in \mathbb{R}^n$. However, depending on the dimensionality of Γ it is not necessarily optimal to include all dimensions. To reduce the number of effective dimensions, we perform a **principal component analysis** (PCA) [70], which is an orthogonal linear transformation returning the eigenvectors (principal components) and eigenvalues of the covariance matrix $\mathbf{\Sigma}$ of the centroid positions. The entries of the covariance matrix

$$\Sigma_{ij} = \mathrm{cov}\left(\mathbf{R}^{(i)}, \mathbf{R}^{(j)} \right) \tag{2.19}$$

are defined as the covariance between the ith and jth component of all centroid positions. The principal components and their corresponding eigenvalues are computed by solving the generalized eigenvalue problem, i.e.,

$$\mathbf{\Sigma} \boldsymbol{M} = \mathbf{\Sigma} \boldsymbol{M} \boldsymbol{D}, \tag{2.20}$$

where $\boldsymbol{M}$ denotes the orthogonal eigenvector matrix and $\boldsymbol{D}$ the diagonal eigenvalue matrix. The principal component $M^{(1)}$ corresponding to the largest eigenvalue coincides with the direction in configuration space exhibiting the largest variance.

Instead of using the original centroid positions to compute cycle centers and diameters, we transform the centroid positions employing the PCA, i.e.,

$$\tilde{\mathbf{R}}_k = \boldsymbol{M}\mathbf{R}_k. \tag{2.21}$$

We expect these transformed centroids to be better suited for describing the cycle centers and diameters as the largest distances in configuration space that can be spanned by a single cycle are more likely to be aligned to the largest principal components.

Together with cycle centers and diameters, cycle affinities, see eq. (2.7), form the **cycle space**. For the above example, the cycle space is spanned by five dimensions: $\boldsymbol{c}_\alpha \in \mathbb{R}^2$, $\boldsymbol{d}_\alpha \in \mathbb{R}^2$ and $A_\alpha \in \mathbb{R}$. Several two dimensional projections of the cycle space are shown in Figure 2.4, in which every point represents a single cycle. These points are clearly not random. In particular, when considering the projection $(c^{(1)}, d^{(1)})$ three well separated point clouds emerge, implying that many different cycles (blue and green) have a small diameter $d_\alpha^{(1)}$ but different cycle centers $c_\alpha^{(1)}$, while the third group of cycles (red) is characterized by large cycle diameters $d_\alpha^{(1)}$. Similar structures are recovered when considering the projections ($c^{(1)}$, cycle affinity), $(c^{(1)}, d^{(2)})$ and $(c^{(2)}, c^{(1)})$. Solely the projections $(c^{(2)}, d^{(2)})$ and ($c^{(2)}$, cycle affinities) exhibit only two distinguishable point clouds.

We can turn these insights into a more quantitative statement by partitioning the cycle space into k communities. To this end we employ an implementation of the **fuzzy c-means** algorithm, which is analogous to the k-means algorithm [cf. section 1.2.1] but assigns to each cycle a probability for belonging to a specific community. First, we normalize all cycle space observables (cycle diameters, cycle centers, and cycle affinities) by their variance to make them comparable. More precisely, we normalize all cycle center and diameter dimensions by the variance of their largest (first) dimension, respectively. For example: $\tilde{c}^{(1)} = c^{(1)} / \operatorname{var}\left(c^{(1)}\right)$ and $\tilde{c}^{(2)} = c^{(2)} / \operatorname{var}\left(c^{(1)}\right)$. The advantage is that the relative variance (weight) between the same type of observable is preserved.

These normalized cycle space observables are then used as an input for the fuzzy c-means clustering algorithm returning membership degrees u_{ij} which express the probability that observation i belongs to community j. To obtain an indicator of how good the clustering results are we compute the **fuzzy partition coefficient** (FPC) that is defined as the Frobenius norm of the membership matrix

$$\mathrm{FPC} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n} u_{ij}^2. \tag{2.22}$$

Here k is the number of chosen communities and n the number of observations (cycles in our

case). Note that other definitions of the FPC exist [146]. In any case, the closer the FPC gets to one the better the cycle space can be partitioned into the chosen number of clusters. The advantage of fuzzy partitioning is that some cycles might not belong to any cycle community while others match well in multiple ones.

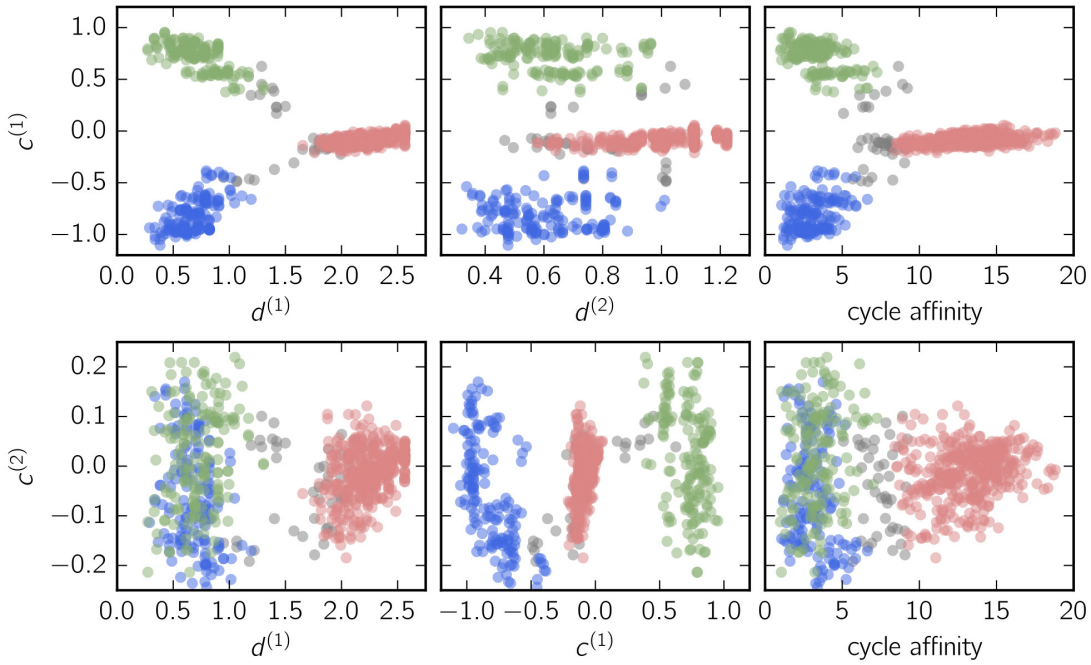

Figure 2.4 – **Scatter plot for several cycle space projections.** Scattered points indicate cycles, while colors (red, blue and green) indicate cycle communities. Gray points indicate cycles that are not assigned to any cycle community.

For the cycle space shown in Figure 2.4 we find $k = 3$ to express the highest FPC, see Figure 2.3(b), which is therefore the best fitting number of cycle communities. The cycle communities detected by the c-means algorithm are highlighted in Figure 2.4 in green, blue (local cycles) and red (global cycles), while gray colored points belong to cycles that are not assigned to any cycle community.

Figure 2.5 depicts the cycle communities projected back onto the (x, y) space. The blue and green colored communities are in agreement with the loci of highest probability, while the red community includes cycles connecting these two loci. This result agrees well with the intuitive picture of entropy production due to the interplay between conservative and nonconservative forces for the particle trapped in a basin, with rare transitions between both basins.

Furthermore, when linking the cycle communities identified in cycle space with their real space projection, we distinguish two different types of cycles: **local cycles** constituting either the blue or green cycle community and **global cycles** constituting the red community. Local cycles are characterized by small cycle diameters and affinities, and similar cycle centers indicating that their corresponding centroids share similar components $\left(\tilde{\mathbf{R}}^{(1)}, \tilde{\mathbf{R}}^{(2)}\right)$ and therefore

belong to a compact region in configuration space. Global cycles, on the other hand, have large cycle diameters and affinities, and connect regions in configuration space that exhibit local cycles.

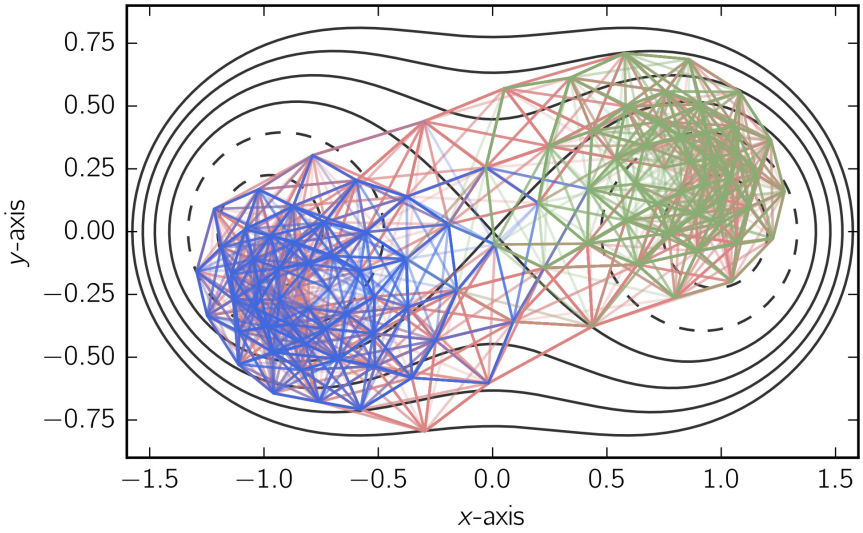

Figure 2.5 – **Cycle communities in configuration space.** Nontrivial cycles are plotted in the $x-y$ plane and colored according to their associated cycle community.

2.5 Community representative and coarse graining

Having grouped cycles into cycle communities, our next goal is to determine a coarse-grained rate matrix $\boldsymbol{W}_{\mathrm{cg}}$ containing as few states as possible but still capturing the dominant (cyclic) dynamics. Unfortunately, it is not possible to compute something like an average cycle because the cycle space does not have any physical metric. For example, we cannot relate the averaged position of, say, the red cycle community in cycle space to a cycle in real space (centroids coordinates).

To overcome this problem, we determine one cycle out of each cycle community, which we will refer to as the **community representative**, and discard all other cycles. Thus, the coarse-grained transition matrix includes states with known centroid coordinates $\mathbf{R}_k$. The final coarse graining step is then to rescale the remaining transition rates.

In the previous section, we discussed how to identify the dissipated heat and total entropy production per transition. Because both quantities solely depend on centroid positions and are time independent, a thermodynamically consistent model ought to preserve them. Additionally, we require that the coarse-grained model preserves the entropy production rate of each cycle community

$$\mathtt{S}_l \equiv \sum_{\alpha\in\mathtt{C}_l} s_\alpha \quad \text{with } \mathtt{C}_l = \Big\{ C_\alpha \;\Big|\; C_\alpha \in l\text{th community} \Big\}. \tag{2.23}$$

When all cycles are uniquely assigned to cycle communities, the total entropy production rate is recovered as the sum over all communities, i.e., $\langle \dot{S}_{\mathrm{tot}} \rangle = \sum_l \mathtt{S}_l$. Persevering the community entropy production rates implies the preservation of the major dynamical properties as well as macroscopic transport. Recently, it has been shown [13, 52] that the fluctuations of current variables (transported quantities), given by $(\delta X)^2 \equiv \langle (X - \langle X \rangle)^2 \rangle / \langle X \rangle^2$, are bounded from above by the average heat dissipation of the process, i.e.,

$$\langle \dot{q} \rangle t = \frac{\langle \dot{S}_{\mathrm{tot}} \rangle}{T} t \geq \frac{2}{(\delta X)^2}, \tag{2.24}$$

where t is the duration of the process. In other words, preserving the total entropy production rate, preserves the upper bound for fluctuations of transported quantities.

2.5.1 Rescaling algorithm

Assume for the moment that appropriate (communities) representatives are known. An algorithm for rescaling the transition rates W_j^i and fluxes Φ_j^i have to fulfill the following physical constraints:

(i) The community entropy production rates $\mathtt{S}_l$ are preserved.

(ii) The dissipated heat q_j^i produced for nonzero transitions is preserved.

(iii) The total entropy A_j^i produced for nonzero transitions is preserved.

(iv) The cycle affinities A_α of the representatives are preserved.

From eq. (2.7) its easy to check that if (iii) holds, then (iv) holds too. In what follows, all rescaled quantities are labeled by hats, e.g., $W_j^i \to \hat{W}_j^i$. Moreover, the subscript l of cycle quantities denotes the cycle representative of the lth community, e.g., s_l is the cycle entropy production rate.

As each cycle community is represented by one cycle only, condition (i) requires that $\mathtt{S}_l$ becomes the representative's entropy production rate, and together with condition (iv) we find the rescaled cycle weight

$$\mathtt{S}_l = \hat{s}_l = \hat{\varphi}_l A_l \quad \Rightarrow \quad \hat{\varphi}_l = \frac{\mathtt{S}_l}{A_l} > 0. \tag{2.25}$$

The crucial coarse graining step consists of removing all other nontrivial cycles by setting their weights $\hat{\varphi}_\alpha$ to zero. All states that are not part of a community representative are thus removed with the remaining states $\hat{V}$ constituting the coarse-grained NE-MSM, see Figure 2.6.

All remaining cycles in the coarse-grained model are either community representatives or trivial. We now show that the weights of the remaining trivial cycles are always positive and

the coarse-grained model still constitutes a valid cycle-flux decomposition

$$\hat{\Phi}^i_j = \sum_\alpha \hat{\varphi}_\alpha \chi^i_{j,\alpha} \tag{2.26}$$

for the rescaled fluxes $\hat{\Phi}^i_j$. By virtue of eq. (2.26) and condition (iii), the new cycle weights $\hat{\varphi}_\alpha$ can then be computed from

$$\exp\left(A^i_j\right) = \frac{\Phi^i_j}{\Phi^j_i} \overset{!}{=} \frac{\hat{\Phi}^i_j}{\hat{\Phi}^j_i} = \frac{\sum_\alpha \hat{\varphi}_\alpha \chi^i_{j,\alpha}}{\sum_\alpha \hat{\varphi}_\alpha \chi^j_{i,\alpha}}, \tag{2.27}$$

which can be rearranged to

$$0 = \sum_\alpha \hat{\varphi}_\alpha \left[\chi^j_{i,\alpha} \exp\left(A^i_j/k_B\right) - \chi^i_{j,\alpha}\right]. \tag{2.28}$$

We now pick one edge $i \to j$ for which the edge affinity $A^i_j > 0$ is positive. We then split the sum over all cycles into a sum over trivial and nontrivial cycles,

$$\begin{aligned} 0 &= \hat{\varphi}_\beta \left[\exp\left(A^i_j/k_B\right) - 1\right] + \sum_{\alpha\neq\beta} \hat{\varphi}_\alpha \left[\chi^j_{i,\alpha} \exp\left(A^i_j/k_B\right) - \chi^i_{j,\alpha}\right] && (2.29)\\ &= \hat{\varphi}_\beta \left[\exp\left(A^i_j/k_B\right) - 1\right] - \sum_l \hat{\varphi}_l \chi^i_{j,l} && (2.30)\\ \Rightarrow \hat{\varphi}_\beta &= \frac{\sum_l \hat{\varphi}_l \chi^i_{j,l}}{\exp\left(A^i_j/k_B\right) - 1}. && (2.31) \end{aligned}$$

For every edge there is exactly one trivial cycle, here denoted β, for which $\chi^i_{j,\beta} = \chi^j_{i,\beta} = 1$. Because the chosen edge affinity A^i_j is positive, all nontrivial cycles are oriented to follow the net probability flow, which implies $\chi^j_{i,\alpha} = 0$. The remaining sum over all nontrivial cycles sharing the edge $i \to j$ reduces to the weight of the community representative since we have set the weight of all other nontrivial cycles to zero. We have thus determined the remaining weights $\varphi_\beta > 0$ of the trivial cycles, which clearly are positive.

Knowing the coarse-grained fluxes $\hat{\mathbf{\Phi}}$, the final step is to determine the coarse-grained transition rates. To fulfill condition (ii), we show – exploiting condition (iii) and inserting the definition of edge affinities [see eq. (1.64)] – that it is sufficient to preserve the ratio of nonzero steady-state probabilities, as

$$\exp\left(\sigma^i_j/k_B\right) = \frac{W^i_j}{W^j_i} = \frac{\pi_j}{\pi_i} \exp\left(A^i_j/k_B\right) = \frac{\hat{\pi}_j}{\hat{\pi}_i} \exp\left(A^i_j/k_B\right) = \frac{\hat{W}^i_j}{\hat{W}^j_i}. \tag{2.32}$$

Therefore, we determine the rescaled steady-state probabilities $\hat{\pi}_i$ by scaling out the probabilities of the removed states through $\pi_i/\pi_j = \hat{\pi}_i/\hat{\pi}_j$ with normalization $\sum_{i\in\hat{V}} \hat{\pi}_i = 1$. Finally,

obtaining the new transition rates $\hat{W}_j^i = \hat{\Phi}_j^i / \hat{\pi}_i$ completes the rescaling algorithm.

2.5.2 Community representatives

Since we know now how to coarse grain a given set of community representatives, we address the question of how to select "appropriate" representatives. Any set of representatives is thought to be appropriate if the graph spanned by their coarse-grained transition matrix is connected and the mean first passage times (MFPTs) between states that belong to local communities are preserved. Preserving MFPTs ensures that the important time scales of the dynamics are preserved too, assuming that these time scales are associated with slow transitions between metastable sets. Instead of identifying metastable sets by means of local communities, however, one can employ a kinetic clustering scheme using hitting times [116], which is in analogy to the often used PCCA+ algorithm [see section 1.3] but also valid for transition matrices breaking detailed balance. For the example shown in Figure 2.4 and 2.5, we identify all states comprised by the blue community as B and by the green community as G. An appropriate set of representatives is required to preserve $\mathrm{MFPT}_{G\to B}$ and $\mathrm{MFPT}_{B\to G}$.

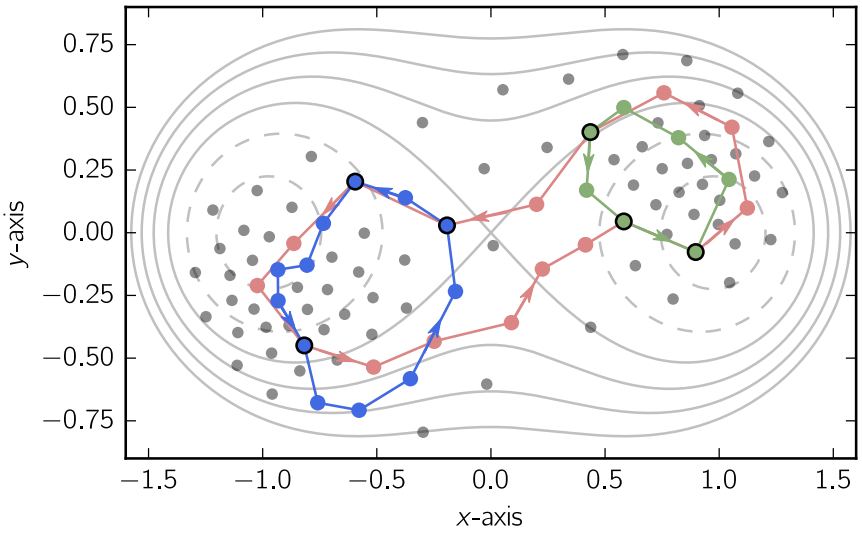

Figure 2.6 – **Illustration of the cycle representatives for each community of the two dimensional model system.** Centroids belonging to community representative are colored according to their respective community, as shown in Figure 2.4, while black edged states highlight intersections of representatives. Gray points indicate centroids not belonging to any representative, which are thus absent in the coarse-grained model.

Next, we introduce a deterministic and a stochastic algorithm for proposing community representatives.

Deterministic

For the deterministic algorithm we define a new graph, the **cycle graph** $\mathscr{G}(V_C, E_C, \Theta)$, where vertices V_C represent cycles and edges E_C links between cycles, while edge weights $\Theta_{\alpha\beta}$ specify the connectivity strength between the αth and βth cycle.

Cycles are thought to be connected if they share a physical state, i.e., at least one vertex on the original graph $\mathscr{G}(V, E, W)$. To quantify their connectivity strength, we define

$$\Theta_{\alpha\beta} \equiv \frac{2\sum_i \pi_i \chi_\alpha^i \chi_\beta^i}{\sum_i \pi_i \chi_\alpha^i + \sum_i \pi_i \chi_\beta^i}, \tag{2.33}$$

which is symmetric as emphasized by the subscripts. The connectivity strength $0 \leq \Theta_{\alpha\beta} \leq 1$ depends on the cumulated probability of the shared states. Other proposals for linking cycles can be found in refs. [7, 33].

The basic idea of the deterministic algorithm is to make use of the concept of the **modularity function** Q which measures the link density inside communities as compared to other communities by assigning it a value between -1 and 1 [94]. The modularity function is defined as

$$Q \equiv \frac{1}{2m}\sum_{\alpha\beta}\left(\Theta_{\alpha\beta} - \frac{k_\alpha k_\beta}{2m}\right) \times \begin{cases} 1 & \text{if } \alpha \text{ and } \beta \text{ belong to same community} \\ 0 & \text{otherwise} \end{cases} \tag{2.34}$$

with $k_\beta \equiv \sum_\alpha \Theta_{\alpha\beta}$ and $m \equiv \frac{1}{2}\sum_\beta k_\beta$.

To make progress, we rewrite the modularity function as the sum over all community modularities

$$Q = \sum_l Q_l = \sum_l \frac{1}{2m}\left(\sum_{\alpha\beta\in C_l} \Theta_{\alpha\beta}^{(l)} - \frac{k_\alpha^{(l)} k_\beta^{(l)}}{2m}\right), \tag{2.35}$$

where l denotes the lth community. Our task is to find the cycle γ for each cycle community l that maximizes $Q_l - Q_{l\setminus\gamma}$, with $Q_{l\setminus\gamma}$ being the modularity of the lth community without cycle γ. In other words, we want to identify the cycle γ for each community that has the biggest impact on the community modularity. In particular, $Q_{l\setminus\gamma}$ increases if cycle γ matches poorly and decreases when it provides many links to other cycles inside its community. If a cycle γ is removed from its community, the difference in its modularity function is given by

$$\begin{aligned} Q_l - Q_{l\setminus\gamma} = \tfrac{1}{2m}\Bigg(& \sum_{\alpha\beta\in C_l\setminus\gamma}\left[\Theta_{\alpha\beta} - \frac{k_\alpha k_\beta}{2m}\right] + \sum_{\beta\in C_l}\left[\Theta_{\gamma\beta} - \frac{k_\gamma k_\beta}{2m}\right] \\ & + \sum_{\alpha\in C_l}\left[\Theta_{\alpha\gamma} - \frac{k_\alpha k_\gamma}{2m}\right] + \frac{k_\gamma^2}{m} - 2\Theta_{\gamma\gamma} - \sum_{\alpha\beta\in C_l\setminus\gamma}\left[\Theta_{\alpha\beta} - \frac{k_\alpha k_\beta}{2m}\right]\Bigg). \end{aligned} \tag{2.36}$$

With $\Theta_{\alpha\beta} = \Theta_{\beta\alpha}$ the difference can be simplified to

$$Q_l - Q_{l\setminus\gamma} = \frac{1}{m}\left(\sum_{\alpha\in C_l}\Theta_{\alpha\gamma} + \frac{k_\gamma\left[k_\gamma - \sum_{\alpha\in C_l} k_\alpha\right]}{2m} - \Theta_{\gamma\gamma}\right). \tag{2.37}$$

Finally, the set of representatives is determined by maximizing eq. (2.37), i.e.,

$$\underset{\gamma}{\operatorname{argmax}}\left(Q_l - Q_{l\setminus\gamma}\right) \tag{2.38}$$

for each community l.

At this point, we want to emphasize that the set of representatives is not necessary connected nor does it need to preserve the slow timescales of the dynamics. If this is the case the next best set of representatives should be examined. Thus the deterministic algorithm should rather be seen as a first guess for an appropriate set of representatives.

Stochastic

Because the deterministic algorithm is not guaranteed to return an appropriate set of representatives, we formulate a stochastic approach which can be easily combined with the deterministic algorithm. The stochastic approach selects randomly candidates for all communities and checks for connectivity and MFPTs. The algorithm is outlined as follows:

(1) Pick one representative per cycle community by drawing a random number.

(2) Check if the set of representatives span an ergodic (connected) transition network. If yes, rescale all transition rates, as described in section 2.5.1, else go back to step (1).

(3) Compute $\mathrm{MFPTs}^{\mathrm{cg}}$ of the coarse-grained model and compare to $\mathrm{MFPTs}^{\mathrm{full}}$ of the full model. If

$$\left|\frac{\mathrm{MFPTs}^{\mathrm{cg}} - \mathrm{MFPTs}^{\mathrm{full}}}{\mathrm{MFPTs}^{\mathrm{full}}}\right| \leq \text{threshold}, \tag{2.39}$$

return the coarse-grained NE-MSM, else go back to step (1).

Step (2) can be efficiently computed by employing the breadth-first search algorithm. If the number of edges of the tree structure returned by the breadth-first search algorithm is less than $|V|-1$, with $|V|$ being the number of coarse-grained states, the transition network is not connected.

Instead of selecting community representatives from the full list of community members, it can be more efficient to combine the stochastic approach with the deterministic algorithm. Therefore, first a list of community representatives, according to the deterministic approach, is determined, e.g., containing 50 nontrivial cycles per community. Then, the stochastic

approach is employed proposing community representatives from this predetermined list only.

3 Polymer in shear flow

So far, we have established a coarse graining procedure for rate matrices breaking detailed balance. From the intuitive example presented in the previous chapter, we understand that probability cycles appear through the interplay between the driving force and the potential energy landscape of the system. However, the discussed example system is somewhat trivial as its potential energy landscape offers well separated energy basins. For more complex systems, the corresponding higher-dimensional potential energy landscape exhibits typically many local minima, which might complicate the identification of cycle communities.

To test our coarse graining procedure, we construct and coarse grain NE-MSMs for a more complex physical model: a tethered polymer under shear.

The rheology of dilute flexible polymers has been studied extensively due to their fundamental and practical relevance [81]. Examples include biomolecules such as the von Willebrand factor in blood plasma and DNA in steady shear flow [127]. The shear drag can overcome the entropic forces favoring coiled or globular configurations and stretch the polymers, which might be a continuous or even discontinuous transition [37]. Motion of DNA tethered to a planar surface has been described as cyclic in experiments [40] and computer simulations [39, 152].

3.1 Polymer model

We study a single model polymer with $N = 50$ beads moving in shear flow close to a substrate. The specific model parameters are inspired by ref. [4]. We employ overdamped Langevin dynamics, cf. section 1.1.2, with

$$\dot{\mathbf{q}}_k = -\nabla_k U + \mathbf{u}_k(\mathbf{q}_k) + \boldsymbol{\eta}_k(t) \tag{3.1}$$

for the bead positions $\mathbf{q}_k = (x_k, y_k, z_k)^T$, where $\mathbf{u}(\mathbf{q})$ represents the shear flow. Interactions with solvent particles are modeled by a random force with correlations $\langle \eta_k^\alpha(t)\eta_l^\beta(t')\rangle = \sqrt{2}\,\delta_{kl}\,\delta^{\alpha\beta}\,\delta(t-t')$, where upper indices label spatial components. The potential energy $U = U_{\mathrm{nb}} + U_{\mathrm{b}} + U_{\mathrm{wall}}$ is split into the nonbonded short-ranged Lennard-Jones pair poten-

tials $U_{\mathrm{nb}} = \varepsilon_{\mathrm{LJ}} \sum_{k<l} \left[r_{k,l}^{-12} - 2 r_{k,l}^{-6} \right]$, bonded interactions $U_{\mathrm{b}} = \frac{\kappa}{2} \sum_{k=1}^{N-1} \left(r_{k+1,k} - 1 \right)^2$ that connect the nearest neighboring beads, and repulsive short-ranged interactions of the wall $U_{\mathrm{wall}} = \varepsilon_{\mathrm{wall}} \sum_{k=1}^{N} z_k^{-12}$. Here $r_{k,l}$ is the distance between the kth and lth bead, $\varepsilon_{\mathrm{LJ}} = 2.3$ determines the strength of the nonbonded potential, $\kappa = 100$ is the effective spring constant, z_k the z-component of the kth bead and $\varepsilon_{\mathrm{wall}} = 2$ the strength of the wall repulsion. The polymer is grafted onto the planar surface (the x-y plane with $z = 0$) by fixing the position of the first bead to $\mathbf{q}_1 = (0, 0, 0.5)$. All quantities have been nondimensionalized by rescaling lengths with the bead diameter σ_{bead} and time scales with the characteristic monomer diffusion time $\sigma_{\mathrm{bead}}^2/(4D)$. Numerical values for the strain rate $\dot{\gamma}$ thus correspond to the Weissenberg number. The polymer is driven into a nonequilibrium steady state through simple shear flow. While some scaling relations depend on hydrodynamic interactions [4, 148], the qualitative behavior of the cyclic motion does not and in the following we neglect hydrodynamic interactions. As flow profile we choose

$$\mathbf{u}_k(\mathbf{q}_k) = \dot{\gamma}(z - z_c)\mathbf{e}_x, \tag{3.2}$$

where $\dot{\gamma}$ is the strain rate and z_c is the z component of the center of mass of the polymer. We found that this shift of the flow stabilizes the globular and stretched configurations as it increases the effective barrier for folding/unfolding and thus leads to a better separation between globular and extended states. Qualitatively, the same effect would be expected when including hydrodynamic interactions with the wall. Our conclusions do not depend on this detail. Although simplified, this model reproduces the cyclic dynamics found in experiments [40]. A schematic illustration of the simulation setup and the cyclic motion of the polymer is shown in Figure 3.1(a).

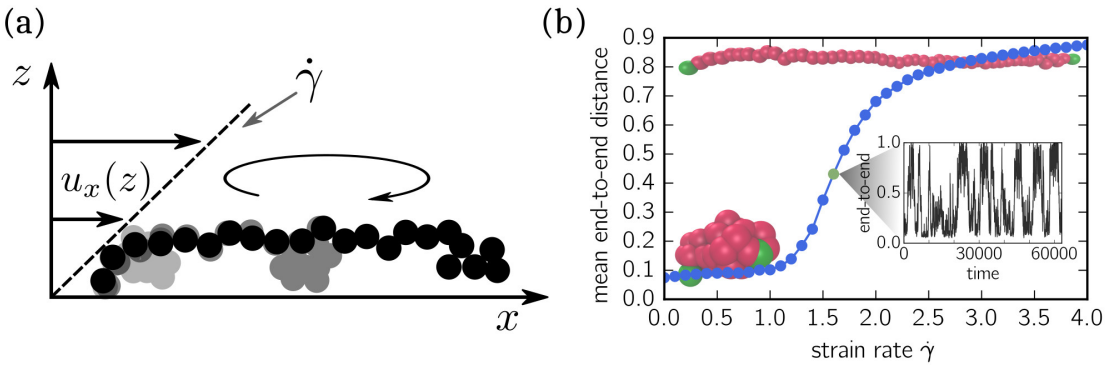

Figure 3.1 – **Polymer in shear flow.** (a) Illustration of polymer dynamics in shear flow. Due to the shear flow $u_x(z)$, the polymer features cyclic folding and unfolding events. (b) Mean end-to-end distance $\langle R_{\mathrm{ee}} \rangle$ as a function of strain rate $\dot{\gamma}$. For $\dot{\gamma} = 1.6$ (green circle) the inset shows an exemplary time series. Also shown are two snapshots for the globule (bottom) and extended polymer (top).

The solvent flow drives the polymer away from thermal equilibrium, which is reflected by the

nonvanishing entropy production rate [125]

$$\langle \dot{S}_{\mathrm{tot}} \rangle = \sum_{k=1}^{N} \left\langle \mathbf{u}_k(\mathbf{q}_k) \cdot \nabla_k U \right\rangle = \dot{\gamma} \langle \sigma_{xz} \rangle \tag{3.3}$$

with off-diagonal stress σ_{xz} of the polymer.

A reasonable order parameter describing the folding and unfolding of the polymer is the relative end-to-end distance

$$R_{\mathrm{ee}} \equiv \frac{1}{N} \left| x_0 - x_N \right|, \tag{3.4}$$

where $R_{\mathrm{ee}} = 1$ corresponds to a straight line of touching beads. We perform computer simulations for multiple values of $\dot{\gamma}$, see Fig. 3.1(b). We find different behaviors of $\langle R_{\mathrm{ee}} \rangle$ that we categorize into three regimes. For $\dot{\gamma} \lesssim 1$ the polymer remains collapsed, while for $\dot{\gamma} \gtrsim 2.2$ it is dominantly found in elongated conformations. For intermediate strain rates the polymer exhibits transitions between globular and elongated conformations, which was also found in similar simulations for free polymers [4] and grafted polymers [82, 39] under shear. The exemplary time series for $\dot{\gamma} = 1.6$ in the inset of Fig. 3.1(b) shows a clear separation of both states with random lifetimes and fast transitions.

3.2 Nonequilibrium Markov state model of the polymer dynamics

Our goal is to construct and coarse grain a dynamically consistent NE-MSM following the procedure introduced in the previous chapter. To capture as many folding and unfolding events as possible we choose in the following the strain rate $\dot{\gamma} = 1.6$.

We start by discretizing the configuration space, which is given by the full set of coordinates of all beads (x_k, y_k, z_k) yielding a total of $3N = 150$ dimensions, employing the k-means clustering algorithm. The configuration space is discretized by $k = 500$ centroids with centroid positions $\mathbf{R}_k$. From the count matrix, we estimate the transition matrix $\boldsymbol{T}(\tau = 1)$ according to a lag time analysis (not shown), and convert it to a rate matrix $\boldsymbol{W}$ [see eq. (1.51)]. Knowing the rate matrix, we determine the stationary distribution π and probability fluxes, e.g., $\Phi^i_j = W^i_j \pi_i$.

Next, we apply the cycle decomposition [cf. section 2.2.1] returning 5192 nontrivial cycles and their cycle weights, together recovering 95% of the total entropy production rate.

To make progress, we apply a PCA to reduce the centroid's dimensionality. Already the first principal component $M^{(1)}$ recovers 94% of the variance of the input data, i.e.,

$$\frac{\lambda^{(1)}_{\mathrm{PCA}}}{\sum_i \lambda^{(i)}_{\mathrm{PCA}}} \leq 0.94, \tag{3.5}$$

where $\lambda^{(1)}_{\mathrm{PCA}}$ is the eigenvalue corresponding to the first principal component.

Similar to the previous chapter, we project the centroid coordinates onto the principal components

$$\tilde{\mathbf{R}}_k = \boldsymbol{M}\mathbf{R}_k, \tag{3.6}$$

where we consider only the two largest components, effectively reducing the 150-dimensional centroid positions $\mathbf{R}_k$ to two-dimensional positions $\tilde{\mathbf{R}}_k$. Figure 3.2(a) depicts the projected centroids, while their colors indicate the end-to-end distance [eq. (3.4)], illustrating the similarity between $M^{(1)}$ and the end-to-end distance.

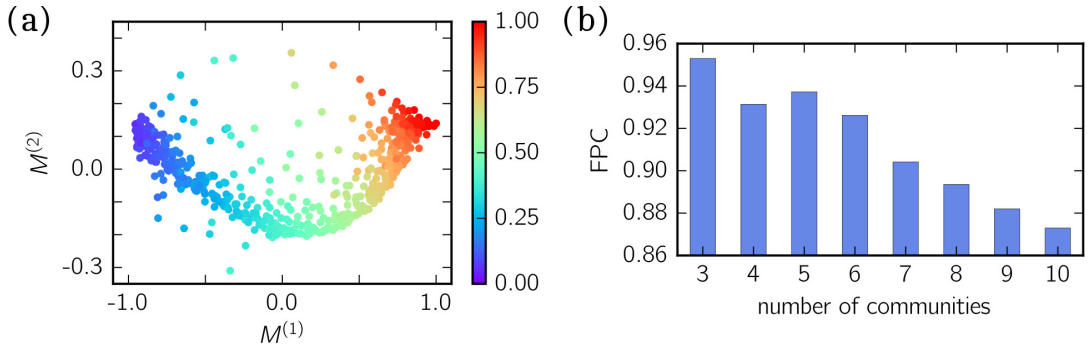

Figure 3.2 – **Configuration space ($\dot{\gamma} = 1.6$) and fuzzy partition coefficient.** (a) Scatter plot of the normalized projections of all centroids onto the two largest principal components $(M^{(1)}, M^{(2)})$. Colors indicate the end-to-end distance R_{ee}. (b) The fuzzy partition coefficient (FPC) computed for multiple cycle communities. The best result is obtained for three communities followed by a slightly lower value for five communities.

Centroids with small (negative) values for $M^{(1)}$ correspond to globular configurations (they contribute little to the observed variance of positions), while large (positive) values correspond to stretched configurations. The second component indicates the variance within these states. Both globular and stretched configurations show larger fluctuations, whereas the intermediate states with $M^{(1)} \sim 0$ exhibit less fluctuations. Hence, the PCA reproduces the expected, typical picture of two basins with intermediate transition states.

Employing the definition of the cycle center, diameter and affinity, see eqs. (2.17),(2.18) and (2.7), every nontrivial cycle is described by the tuple $(c_\alpha^{(1)}, d_\alpha^{(1)}, A_\alpha)$ spanning the cycle space. We discard the remaining dimensions of the cycle center and diameter $(c^{(2)}, d^{(2)})$, because their absolute values are negligible in comparison with $(c^{(1)}, d^{(1)})$.

3.2.1 Cycle space

The cycle space is illustrated in Figure 3.3, clearly showing that cycles are not distributed randomly. In more detail, many cycles have a small diameter $d_\alpha^{(1)}$ but many different cycle centers $c^{(1)}$. Following the nomenclature of the previous chapter, see section 2.4, these cycles are identified as local cycles. The group of cycles exhibiting larger diameters are accordingly

identified as global cycles. In principal, the same classification holds when taking the cycle affinities into account, i.e., cycles exhibiting large cycle affinities are identified as global cycles, while cycles with low cycle affinities are identified as local cycles.

The next step is to group cycles together forming cycle communities for which we normalize all cycle space dimensions by their variance and employ the c-means algorithm [cf. section 2.4]. Figure 3.2(b) depicts the fuzzy partition coefficient (FPC) for a different number of communities. The largest FPC is found for $k = 3$ cycle communities followed by a slightly lower value for $k = 5$.

To illustrate the difference in the detected cycle communities, Figure 3.3(a) and (b) displays the cycle space for three and five detected communities, respectively. For $k = 3$ the c-means algorithm detects two cycle communities (blue and green) including local cycles and one cycle community (red) including global cycles. Moreover, the cycles of the blue community contain centroids associated with globule (coiled) conformations, while the green community is associated with stretched configurations. The cycles of the red community are connecting both regions in configuration space and thus are responsible for the folding and unfolding transitions.

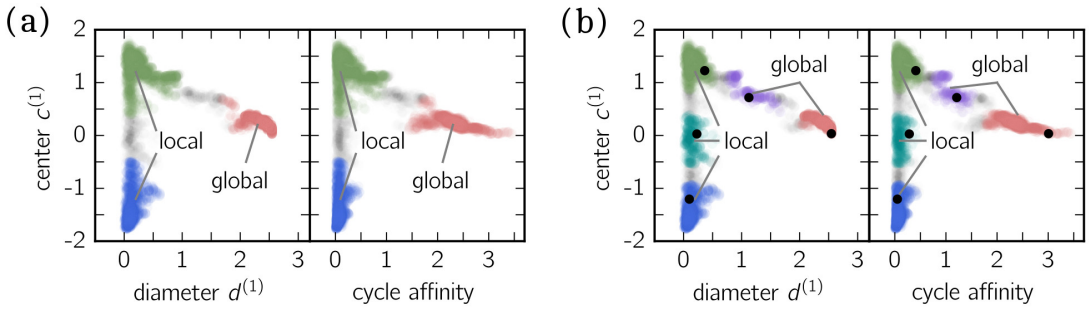

Figure 3.3 – **Cycle space ($\dot{\gamma} = 1.6$).** (a) Scatter plot of cycle center vs. the cycle diameter, and cycle affinity vs. cycle center for $k = 3$ cycle communities. Colors represent cycle communities, while gray points indicate no community. (b) The same plots as for (a) but for five cycle communities. Black points indicate the cycle representatives.

A similar picture is obtained when grouping all cycles into five cycle communities. Here the blue, green and red colored communities are basically the same as determined for $k = 3$ cycle communities. The additional purple and cyan communities include global and local cycles, respectively. From Figure 3.2(a) we understand that the cyan colored local cycles with $c^{(1)} \approx 0$ are located at intermediate end-to-end distances ($M^{(1)} \approx 0$) and therefore contain half-stretched configurations. Furthermore, the purple colored global cycles exhibit smaller cycle diameters than the red community but larger cycle centers, suggesting that purple cycles connect half-stretched and fully extended configurations.

Although the minimal number of cycle communities accounting for the collective folding and unfolding dynamics of the polymer is three, we choose $k = 5$ communities in the following,

which allows to capture the dynamics of the polymer in more detail.

3.2.2 Community representatives and coarse graining

In the next step for constructing a coarse-grained NE-MSM, we need to find suitable community representatives, delete all remaining cycles and rescale all nonzero transition rates. Here we employ the stochastic algorithm, see section 2.5.2, returning a coarse-grained NE-MSM including five nontrivial cycles. The stochastic algorithm ensures that the rate matrix $\boldsymbol{W}_{\mathrm{cg}}$ preserves the slow time scales of the system by mapping MFPTS. In particular, all MFPTs between sets of configurations that are part of local communities are preserved. The positions of the determined community representatives in cycle space are highlighted by black dots in Figure 3.3(b).

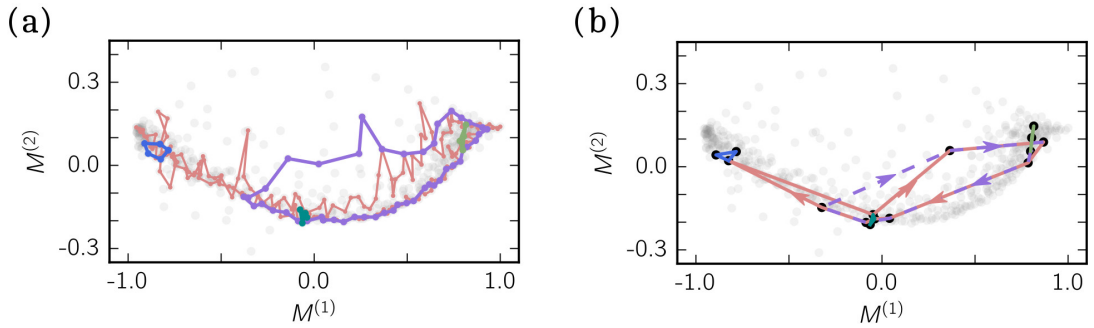

Figure 3.4 – **Transition network in configuration space** using the normalized projections onto the two largest principal components before (a) and after removing all triangle and bridge states (b). Symbols represent centroids, while lines indicate cycles with arrows [(b) only] pointing in the direction of the net flux. Colors match the community colors as shown in Figure 3.3.

At this stage the coarse-grained NE-MSM still contains many states since a single cycle can traverse hundreds of states. The community representatives that form the coarse-grained NE-MSM are depicted in Figure 3.4(a). The important point is, however, that the coarse-grained model has lost much of its original complexity as it now contains only a few cycles. We can thus further reduce the number of states. To this end, we identify two dominant motifs, which we refer to as **bridge and triangle states**. Both motifs include states that have exactly two neighbors.

3.2.3 Coarse graining of triangle and bridge states

Bridge states

To coarse grain bridge states [state 1 in Figure 3.5(a)] we form a new connection between the two neighboring states [state 0 and 2 in Figure 3.5(a)]. Similar to the rescaling of community representatives, the rescaling algorithm for bridge states, more precisely the new transition

rates $\hat{W}_2^0$ and $\hat{W}_0^2$, have to fulfill the following conditions:

(i) The new-formed edge affinity balances the lost edge affinities, i.e., $\hat{A}_2^0 = A_1^0 + A_2^1$.

(ii) No probability current is lost, i.e., $\hat{J}_2^0 = J_1^0 = J_2^1$.

(iii) The produced medium entropy of the new-formed transition balances the lost medium entropy production, i.e., $\log\left(\hat{W}_2^0/\hat{W}_0^2\right) = \log\left(W_1^0/W_0^1\right) + \log\left(W_2^1/W_1^2\right)$.

(iv) Edge affinities and medium entropy production of the remaining part of the network is preserved. i.e., $\hat{A}_j^i = A_j^i$ and $\log\left(\hat{W}_j^i/\hat{W}_i^j\right) = \log\left(W_j^i/W_i^j\right)\ \ \forall i,j \neq \{0,1,2\}$.

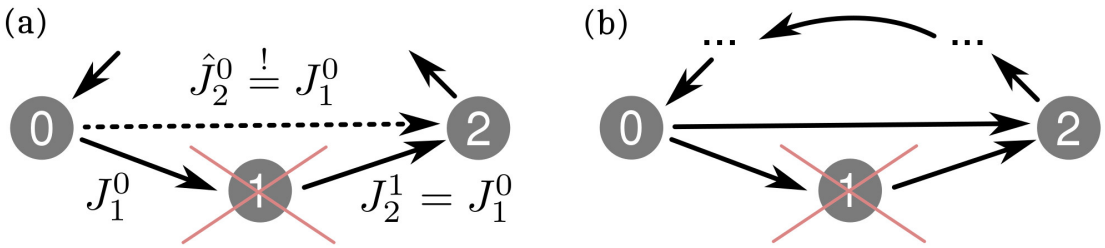

Figure 3.5 – **Illustration of the bridge and triangle state coarse-graining approach.** (a) Bridge state: After deletion of state 1 the other states are directly connected (dotted arrow). (b) Triangle state: After deletion of state 1 the existing connections $0 \leftrightarrow 2$ are modified. The arrows point in the direction of the probability currents.

Making use of condition (ii) and (i), i.e.,

$$\hat{J}_2^0 = J_1^0 = \hat{\Phi}_2^0 - \hat{\Phi}_2^0 = \Phi_1^0 - \Phi_1^0 \quad \text{and} \quad \log\left(\hat{A}_2^0/k_B\right) = \frac{\hat{\Phi}_2^0}{\hat{\Phi}_0^2} = \frac{\Phi_1^0\Phi_2^1}{\Phi_0^1\Phi_1^2},$$

the new probability fluxes follow as

$$\hat{\Phi}_0^2 = \frac{\Phi_1^0 - \Phi_0^1}{\left(\frac{\Phi_1^0\Phi_2^1}{\Phi_0^1\Phi_1^2}\right) - 1} \quad \text{and} \quad \hat{\Phi}_2^0 = \frac{\Phi_1^0 - \Phi_0^1}{\left(\frac{\Phi_0^1\Phi_1^2}{\Phi_1^0\Phi_2^1}\right) - 1}. \tag{3.7}$$

Inserting condition (i) in condition (iii)

$$\frac{\hat{W}_2^0}{\hat{W}_0^2} = \frac{\hat{\Phi}_2^0}{\hat{\Phi}_0^2}\frac{\hat{\pi}_0}{\hat{\pi}_2} \stackrel{!}{=} \frac{W_1^0 W_2^1}{W_0^1 W_1^2} = \frac{\Phi_1^0\Phi_2^1}{\pi_0\pi_1}\frac{\pi_2\pi_1}{\Phi_0^1\Phi_1^2} = \frac{\hat{\Phi}_2^0}{\hat{\Phi}_0^2}\frac{\pi_0}{\pi_2}, \tag{3.8}$$

we find that condition (iii) is fulfilled if the ratio between the two probabilities $\hat{\pi}_0/\hat{\pi}_2 = \pi_0/\pi_2$ is preserved. Rescaling π_0 and π_2, however, requires the rescaling of all other probabilities π_i as otherwise condition (iv) is violated. With the same reasoning as shown for the rescaling algorithm introduced in section 2.5.1, all other probabilities are rescaled by $\hat{\pi}_i = \hat{\pi}_j \frac{\pi_i}{\pi_j}$ with $\sum_i \hat{\pi}_i = 1$.

One should note that condition (iv) is also the main difference between our approach and the one discussed in ref. [8]. While in our adaptation the probability distribution of the complete network is changed, Altaner *et al.* change it only locally (π_0 and π_2) and hence absorb π_1 into π_0 and π_2. The disadvantage of the latter is that condition (iv) is not preserved and, when using the coarse-graining approach iteratively, accumulation of probability in single states might occur, which leads to unphysical results.

Triangle states

For the coarse graining of triangle states [state 1 in Figure. 3.5(b)] we consider all cycles that contain the motif $0 \to 1 \to 2$. We modify these cycles by replacing the edges $0 \to 1 \to 2$ with a new edge $0 \to 2$. To be thermodynamically consistent the modified cycles have to fulfill the following conditions:

(i) All nontrivial cycles and their entropy production rates, i.e., $s_\alpha = \varphi_\alpha A_\alpha$ are preserved.

(ii) All edge affinities A^i_j not involving state 1, in particular A^0_2, are preserved.

(iii) The medium entropy produced for all transitions not involving state 1, i.e., $\sigma^i_j = k_B \log\left(W^i_j / W^i_j\right)$, are preserved.

These conditions are very similar to the conditions formulated for the rescaling algorithm for the community representatives [cf. section 2.5.1] and thus it is not surprising that the same rescaling algorithm can be applied to remove triangle states. Note, however, that through restriction (i) it is not necessarily possible to coarse grain all found triangle structures. Assume, for instance, that the modified cycle coincides with an already existing cycle, then the rescaling is not unique anymore and entropy production is destroyed as only one out of two cycles survives.

Given a system with many bridge and triangle states, as illustrated in Figure 3.4(a), the general approach for removing all possible bridge and triangle states is to first remove iteratively all bridge states, then remove iteratively all possible triangle states and finally remove iteratively all new-formed bridge states. Figure 3.4(b) shows the final coarse-grained transition network projected onto the two largest principal components after removing all possible bridge and triangle states.

3.2.4 Final coarse-grained NE-MSM

The final coarse-grained NE-MSM for strain rate $\dot{\gamma} = 1.6$ is shown in Figure 3.6. After removing the bridge and triangle states the transition network contains 15 states (centroids). The collective rates for folding and unfolding of this coarse-grained model agree with those obtained from the full Markov model by construction. Moreover, the remaining five cycles now allow detailed insights into the relevant pathways in nonequilibrium. The three local cycles are

composed of three states, the minimal number for a nontrivial, entropy-producing cycle. The blue and green cycle connect stable globular and stretched configurations, respectively. The cyan cycle represents a metastable intermediate of half-stretched configurations that do not unfold correctly but quickly fold back to the intermediate. The global purple cycle also contains half-stretched configurations (structurally similar to the cyan colored cycle) but here the unfolding reaches the final stretched states before returning to their half-stretched origin. Finally, the red cycle represents the full transition from globule to stretched configurations and back.

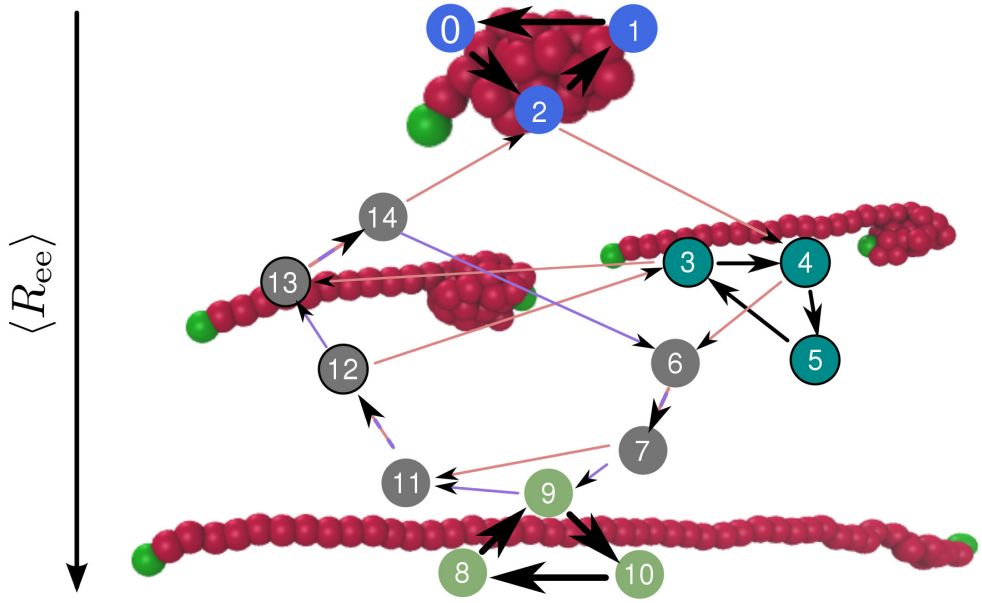

Figure 3.6 – **Final NE-MSM for $\dot{\gamma} = 1.6$.** Transition network of the polymer dynamics with 15 centroids (filled circles) and five cycles. The colored states green, blue, and cyan correspond to the colors of the local cycle communities shown in Figure 3.3(b). States with a black border are structurally very similar and constitute the transition ensemble. The arrows point in the direction of probability currents (net flux), while the arrow widths represent the magnitude of currents. On average, the polymer dynamics follows the direction of the arrows. Arrow colors red and purple correspond to the global cycles.

The five states $(3,4,5,12,13)$ describing intermediate, half-stretched configurations are very close in configuration space, see Figure 3.4(b). They constitute the analog of the transition ensemble through which the folding/unfolding has to proceed. In nonequilibrium, however, the folding and unfolding processes follow different paths through this narrow region of configuration space. The globule to half-stretched transition proceeds along $2 \rightarrow 4$ (with state 4 belonging to the cyan cycle), whereas the reverse half-stretched to globule transition proceeds along $13 \rightarrow 14 \rightarrow 2$, with states 13 and 14 belonging to the purple cycle. The cycle topology thus reveals the dynamical trapping of the polymer in an intermediate, which cannot be captured by structural information alone. Employing order parameters like distances in configuration space or lifetimes to identify mesostates will clearly miss this important feature

of cyclic nonequilibrium dynamics.

Another question that we can address is dissipation, the role of which for biomolecular processes has been investigated recently, e.g., for self-replication [41] and in the activation of signaling proteins [147]. The rate of dissipated heat $\dot{q}_l = \mathrm{S}_l / T$ created in each cycle community is proportional to their respective entropy production rate. Our analysis reveals that both the green and blue cycles are equally responsible for about 30% of the total dissipated heat, while the cyan, purple and red communities produce 5%, 15% and 50%, respectively. The latter is caused by the large conformational changes (folding and unfolding process) of the polymer. The blue and red cycles, on the other hand, do not exhibit large conformational changes, therefore, the conformational changes must be on very short time scales, which is confirmed by the large probability currents shown in Figure 3.6.

4 Building Markov state models for periodically driven nonequilibrium systems

By now, we discussed how to construct nonequilibrium Markov state models for systems driven into a NESS via time-independent driving mechanisms. Another possibility to reach a NESS is, however, to drive a system periodically in time. Time periodic driving offers a variety of applications, for example molecular systems can be probed and even steered by exposing them to oscillating electromagnetic fields [11, 24, 42] or by applying periodically external mechanical forces [22, 91].

In this chapter, we discuss two methods of how to construct NE-MSMs for periodically driven systems. We then apply both methods to two example systems and study their response to different oscillation periods.

4.1 Periodic Driving

We start by assuming that the system of interest is driven out of equilibrium due to a time-dependent oscillating protocol λ with period T, i.e., $\lambda(t) = \lambda(t+T)$. For now we restrict the protocol to act on the dynamics via an explicitly time-dependent potential, i.e., $U(\lambda(t+T)) = U(\lambda(t))$. As a paradigmatic description for the time evolution of the system, we consider overdamped Langevin dynamics or rather its associated Fokker-Planck equation [cf. section 1.1.6], i.e.,

$$\frac{\partial \rho(t)}{\partial t} = \mathscr{W}\,\rho(t) \quad \text{with } \mathscr{W} = \left[D\,\Delta + \left(-\nabla U\big(\lambda(t)\big)/\gamma_{\text{fric}} \right) \cdot \nabla \right]^{\dagger}, \tag{4.1}$$

although the same argumentation holds for thermostated Hamiltonian as well as general Langevin dynamics. Due to the periodicity of λ, the adjoint Fokker-Planck generator and its space discretized rate matrix become periodic in time too, i.e., $\boldsymbol{W}(t) = \boldsymbol{W}(t+T)$. The corresponding master equation reads

$$\frac{\partial p^{\mathrm{T}}(t)}{\partial t} = p^{\mathrm{T}}(t)\,\boldsymbol{W}(t), \tag{4.2}$$

while its solution is given by

$$p^{\mathrm{T}}(t) = p^{\mathrm{T}}(0)\ \boldsymbol{K}(t). \tag{4.3}$$

Here $\boldsymbol{K}(t) \in \mathbb{R}^{N\times N}$ is the fundamental matrix propagator solving

$$\frac{\partial}{\partial t}\boldsymbol{K}(t) = \boldsymbol{K}(t)\ \boldsymbol{W}(t) \tag{4.4}$$

with $\boldsymbol{K}(0) = \mathbf{1}$. For fixed protocol λ, the rate matrix does not depend on time and thus the solution of eq. (4.2) is given by $p^{\mathrm{T}}(t) = p^{\mathrm{T}}(0)\ \exp\big(\boldsymbol{W}t\big)$.

Solving eq. (4.4) for general time-dependent protocols $\lambda(t)$ requires a numerical solution, returning for every point in time another matrix $\boldsymbol{K}(t)$.

4.1.1 Floquet theory

If the protocol is periodic in time [$\lambda(t+T) = \lambda(t)$], however, we can make use of **Floquet's theorem** stating that the solution of eq. (4.4) can be split into a periodic and nonperiodic part [78], which reads

$$\boldsymbol{K}(t) = \boldsymbol{Q}(t)\ \exp\big(\tilde{\boldsymbol{W}}t\big) \quad \text{with} \quad \boldsymbol{Q}(t+T) = \boldsymbol{Q}(t). \tag{4.5}$$

Together with eq. (4.3), the time-dependent probability distribution follows as

$$p^{\mathrm{T}}(t) = p^{\mathrm{T}}(0)\ \boldsymbol{Q}(t)\ \exp\big(\tilde{\boldsymbol{W}}t\big). \tag{4.6}$$

Here the time-independent matrix $\tilde{\boldsymbol{W}}$ is of particular importance as it describes the long-term evolution of the system and is at the center of our subsequent investigation. Because of the probability conserving characteristics of the rate matrix $\boldsymbol{W}(t)$, the matrix $\tilde{\boldsymbol{W}}$ also conserves probabilities and hence has the form of an effective time-independent rate matrix. In the following we introduce two different approaches how to approximate $\tilde{\boldsymbol{W}}$.

4.1.2 Approximating the rate matrix $\tilde{\boldsymbol{W}}$

Direct approach

The first approach, which we refer to as **direct** approach, was first introduced by Wang *et al.* [144]. Here we follow the standard Markov state modeling procedure as introduced in section 1.2. Therefore, we use trajectories from MD simulations [$\lambda(t)$ is incorporated in the dynamics] to sample the configuration space, discretize it, for example by employing k-means clustering, and estimate all relevant transition probabilities and microscopic rates. The only restriction is that the chosen lag time must be in multiples of the period, i.e., $\tau = nT$ with

$n \in \mathbb{N}$. To show this, consider eq. (4.5) for $t = nT$,

$$\boldsymbol{K}(nT) = \underbrace{\boldsymbol{Q}(0)}_{=\boldsymbol{K}(0)=\mathbf{1}} \exp(\tilde{\boldsymbol{W}} nT) = \boldsymbol{K}^n(T). \tag{4.7}$$

Thus, the probability distribution for multiples of T is given as

$$p^{\mathrm{T}}(nT) = p^{\mathrm{T}}(0)\, \boldsymbol{K}^n(T). \tag{4.8}$$

The estimated matrix propagator $\boldsymbol{K}(T)$ can be identified as time-discrete transition matrix [to avoid confusion we keep $\boldsymbol{K}$ instead of $\boldsymbol{T}$], i.e., $K_j^i \geq 0$ and $\sum_j K_j^i = 1$. Employing eq. (1.51), $\boldsymbol{K}(T)$ is converted [see eq. 1.51] to the desired time-independent rate matrix $\tilde{\boldsymbol{W}}_{\mathrm{dir}}$, where the subscript denotes the direct approach.

Indirect approach

For the second approach, which we refer to as **indirect** approach, consider eq. (4.2) where the rate matrix depends on time via the protocol $\lambda(t)$, i.e., $\boldsymbol{W}(t) = \boldsymbol{W}\big(\lambda(t)\big)$. When freezing the protocol at time $t = t_k$, that is, $\lambda(t_k) \equiv \lambda_k$, the solution of eq. (4.2) reaches its steady state distribution for $t \to \infty$, i.e.,

$$\pi^{\mathrm{T}}(\lambda_k)\, \boldsymbol{W}(\lambda_k) = 0. \tag{4.9}$$

Moreover, each $\boldsymbol{W}(\lambda_k)$ fulfills the detailed balance condition $\left[W_j^i(\lambda_k)\, \pi_i(\lambda_k) = W_i^j(\lambda_k)\, \pi_j(\lambda_k)\right]$ as for every fixed value λ_k the dynamics is governed by a potential $U(\lambda_k)$ that does not depend on time anymore. In other words, the rate matrix for every single point in time is governed by equilibrium dynamics obeying detailed balance. Principally, for every λ_k the configuration space of the system can be sampled by employing MD simulations and the microscopic rate matrices $\boldsymbol{W}(\lambda_k)$ can be estimated. In practice, however, when $\lambda(t)$ is a continuous function it has to be approximated by a finite number of λ_k's.

Once a set of matrices $\{\boldsymbol{W}(\lambda_k) \mid k \in \mathbb{N}\}$ is obtained, $\tilde{\boldsymbol{W}}_{\mathrm{indir}}$ can be computed by solving eq. (4.4) for $\boldsymbol{K}(t = T)$ through direct numerical integration and subsequent conversion of $\boldsymbol{K}(T)$.

4.1.3 Remarks

Does $\tilde{\boldsymbol{W}}$ fulfill detailed balance ?

In general, $\tilde{\boldsymbol{W}}$ does not fulfill detailed balance, although the generator of the dynamics is approximated by matrices fulfilling it. Instead, $\tilde{\boldsymbol{W}}$ reaches a NESS causing time-independent probability currents to flow through the system. In the limit of an infinitely slow protocol $\lambda(t)$, detailed balance is recovered as the dynamics is in quasiequilibrium. Then, however, $\tilde{\boldsymbol{W}}$ becomes zero and the change in dynamics is solely governed by $\boldsymbol{Q}(t)$.

Direct versus indirect approach

The advantage of the direct approach is its lower computational cost in comparison with the indirect approach. The protocol $\lambda(t)$ is incorporated in the MD simulations and the NE-MSM can be directly constructed from trajectories. The standard MSM machinery, however, can only be partially applied since detailed balance, the keystone of equilibrium Markov State Modeling, does not hold. Another disadvantage is that, since the shortest lag time for building a NE-MSM is the period T, the dynamics on timescales shorter than T cannot be reliably predicted.

The indirect approach, on the other hand, is computationally more expensive as for every λ_k an individual MSM must be constructed and thus the number of required trajectories grows proportional with the number of needed λ_k's. The single MSM, though, obeys detailed balance and can therefore be constructed by employing well-developed algorithms. However, the largest advantage is that, once $\lambda(t)$ is sufficiently fine discretized and all corresponding MSMs are obtained, new protocols can be constructed that contain the same λ_k's, i.e., the ordering or individual duration, or both can be changed.

Of particular importance could be to change the period T in order to investigate its influence on the dynamics. Within the indirect approach this is easily accomplished as only eq. (4.4) needs to be solved for a new protocol. Note that the indirect approach does not require new MD simulations. For the direct approach, on the contrary, every new protocol requires the computation of new MD trajectories and the construction of new MSMs.

Nonconservative drift term

In eq. (4.1) the protocol $\lambda(t)$ enters directly into the potential energy term. It is straightforward to extend both described approaches to include protocols not acting on the potential energy but on a general drift term. The corresponding Fokker-Planck operator reads

$$\mathscr{W} = \left[D\,\Delta + \Big(-\nabla U/\gamma_{\mathrm{fric}} + \mathbf{f}\big(\lambda(t)\big) \Big) \cdot \nabla \right]^{\dagger} . \tag{4.10}$$

The direct approach does not need any modification and can be applied as it stands. The only consequence for the indirect approach is that the rate matrices $\boldsymbol{W}(\lambda_k)$ do not necessarily fulfill detailed balance, which has to be regarded for their construction.

4.2 Three-state system

We start by giving an intuitive and simple example demonstrating the direct and indirect approach.

Let $\lambda(t)$ be a two-step protocol given by

$$\lambda(t) = \begin{cases} 0, & 0 \le t \le T/2 \\ 1, & T/2 \le t \le T \end{cases} \quad \text{and} \quad \lambda(t+T) = \lambda(t). \tag{4.11}$$

The associated rate matrix $\boldsymbol{W}(t)$ is given as

$$\boldsymbol{W}(t) = \boldsymbol{W}^{(0)} \left[1 - \lambda(t)\right] + \boldsymbol{W}^{(1)} \lambda(t) \tag{4.12}$$

with

$$\boldsymbol{W}^{(0)} = \begin{pmatrix} -1.8 & 0.3 & 0.4 \\ 0.6 & -0.33 & 0.02 \\ 1.2 & 0.03 & -0.42 \end{pmatrix} \quad \text{and} \quad \boldsymbol{W}^{(1)} = \begin{pmatrix} -0.78 & 2.6 & 0.34667 \\ 0.26 & -2.86 & 0.01733 \\ 0.52 & 0.26 & -0.364 \end{pmatrix} \tag{4.13}$$

fulfilling detailed balance and stationary distributions

$$\pi^{(0)} = \begin{pmatrix} 0.1667 & 0.3332 & 0.5 \end{pmatrix}^{\mathrm{T}} \quad \text{and} \quad \pi^{(1)} = \begin{pmatrix} 0.38461 & 0.03846 & 0.57692 \end{pmatrix}^{\mathrm{T}}. \tag{4.14}$$

To illustrate the direct approach we compute the transition matrices $\boldsymbol{K}^{(0)}$ and $\boldsymbol{K}^{(1)}$ with lag time $\tau = 0.01$ from $\boldsymbol{W}^{(0)}$ and $\boldsymbol{W}^{(1)}$, respectively, generate a sufficiently long Markov chain according to the protocol in eq. (4.11) with $T = 0.4$, and construct $\tilde{\boldsymbol{W}}_{\mathrm{dir}}$ estimating the transition matrix and converting it to the desired rate matrix.

In order to estimate the quality of $\tilde{\boldsymbol{W}}_{\mathrm{dir}}$, the time evolution of the probability distribution is computed by numerically integrating the protocol with initial distribution $p(0) = (0\ 0\ 1)^T$. The results are shown in Figure 4.1(a), which clearly shows that the probability distribution $p^{\mathrm{dir}}(t)$ (red curves) follows the averaged oscillating distribution of the full solution. In Figure 4.1(b), the same time evolution is shown but for the indirect approach. Again, the constructed solutions (red curves) follow well the long-term behavior of the oscillating distributions, although their numerical values do not follow the averaged oscillating curves but rather their local extrema. To quantitatively compare the dynamics of the direct/indirect approach, we compute time scales of the respective rate matrices

$$\tau_{\mathrm{dir}} = \{1.060, 0.447\} \quad \text{and} \quad \tau_{\mathrm{ind}} = \{1.031, 0.434\}, \tag{4.15}$$

which are in very good agreement.

Employing the indirect approach, the dynamics of the system for different protocol periods can be examined. Aside from timescales, the current J flowing through the system and the mean entropy production rate $\langle \dot{S} \rangle$ are of special interest. Following the discussion of the previous chapters, probability current can only flow in cycles. In a three-state system only one entropy producing cycle exists which we identify as $C = \{0, 1, 2, 0\}$. While its cycle weight denotes the magnitude of the net flux flowing through the system, the cycle itself dictates the

dynamics of the system, i.e., what state follows a given state most often in time.

Figure 4.1(c) shows the magnitude of the current J flowing through cycle $0 \to 1 \to 2 \to 0$ and mean entropy production rate for different periods T. Interestingly, neither J nor $\langle \dot{S} \rangle$ are monotonic in T but rather show a distinct maximum at almost the same T_{res}, which is known as stochastic resonance [51]. The meaning of T_{res} is that here the system shows the largest magnitude of cyclic motion which coincides with a maximum in entropy production.

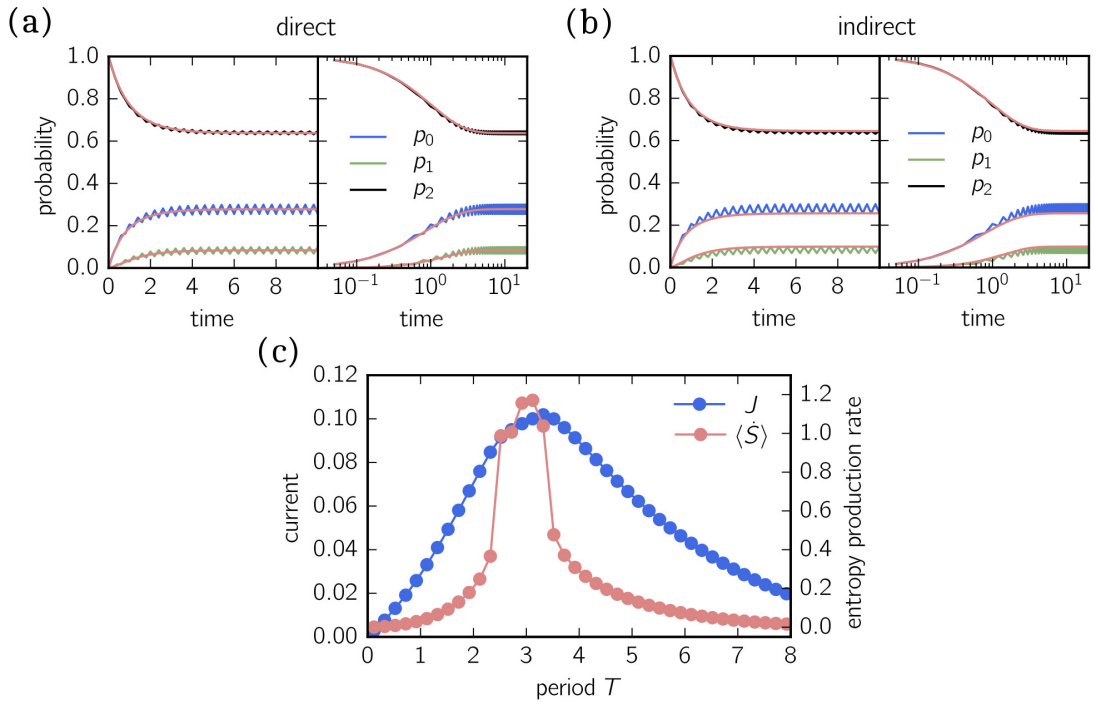

Figure 4.1 – **Direct and indirect approach for the three-state system.** (a+b) Time evolution of all three state probabilities (left panel regular scaling and right panel semi-logarithmic scaling) is shown in blue, green and black, respectively. Probabilities following the direct/indirect approach are colored in red. The initial probability distribution is set to $p(0) = (0\ 0\ 1)^{\mathrm{T}}$. (c) The current flowing through the cycle $0 \to 1 \to 2 \to 0$ and the mean entropy production rate for different protocol periods T.

4.3 Alanine dipeptide in water

As a second example we examine alanine dipeptide in aqueous solution. Alanine dipeptide is a small peptide that serves as a paradigm for computational studies for free energy calculations and rate estimation of proteins [46, 29, 38, 140, 145].

4.3.1 Computational details

To study alanine dipeptide in water we employ the molecular dynamics software Gromacs 5.1.2 [1]. The alanine dipeptide molecule is modeled with the CHARMM22/CMAP [86] force field, dissolved in 362 TIP3P [71] water molecules and positioned in a cubic box (2.25 nm box length) with periodic boundary conditions. Long-range electrostatics are treated using particle mesh Ewald summation [36] with cubic interpolation and Fourier grid spacing of 0.16 nm. The cutoff for all short-ranged interactions is set to 1.0 nm. All hydrogen involving covalent bonds are constraint by the LINCS algorithm [58]. For the time step is set to 2 fs. All simulations are conducted in the NPT ensemble, for which the temperature is set to 300 K using the velocity-rescaling [25] thermostat with $\tau_T = 1$ ps. For isotropic pressure coupling we employ the Parrinello-Rahman [102] barostat with $\tau_p = 2$ ps.

4.3.2 Free energy landscape and driving protocol

One reason that makes alanine dipeptide a popular model systems is that its molecular conformations are well characterized by its two dihedral angles ϕ and ψ, as illustrated in Figure 4.2(a).

To drive the system out of equilibrium, a periodic protocol is employed that has the physical meaning of an oscillating electric field caused, for example, by a laser beam. We assume the electric field to follow a cosine function in time with amplitude A, period T and its direction set parallel to the x-axis. We choose the electric field to point purely in positive x-direction because the (ϕ,ψ) configuration space would otherwise be degenerated with respect to positive/negative field directions, making it impossible to compare both approaches within the (ϕ,ψ) space representation. The external electric field reads

$$\mathbf{E}(t) = E_0 \left(\cos\left(\frac{2\pi t}{T}\right) + 1, 0, 0\right)^{\mathrm{T}} \equiv \lambda(t)\, \mathbf{e}_x, \tag{4.16}$$

with amplitude E_0 set to 0.5 V/nm, which is similar to the study of Wang *et al.* [145] in which it is set to 1 V/nm. To be able to apply the indirect approach, we discretize the continuous protocol $\lambda(t)$ by 20 steps. Both continuous and discrete protocols are shown in Figure 4.2(b). Because of the symmetry of the cosine function, however, the number of equilibrium MSMs that must be constructed is reduced by a factor two. For the direct approach we conduct computer simulations directly incorporating the oscillating electric field with period $T = 5$ ps.

Before comparing both approaches, we discuss the influence of a static electric field on the free energy landscape. In Figure 4.2(c), we show the free energy landscapes [(ϕ,ψ) representation] for three different field strengths. Without an applied electric field we identify three different free energy basins known as the extended conformation of the backbone $\left[C7_{\mathrm{eq}} = \{(\phi,\psi) \mid -180^\circ \le \phi \le 0^\circ \text{ and } 130^\circ \le \phi \le 180^\circ,\ -180^\circ \le \psi \le -100^\circ \text{ and } 75^\circ \le \psi \le 180^\circ\}\right]$, left-handed $\left[\alpha_L = \{(\phi,\psi) \mid -180^\circ \le \phi \le 0^\circ \text{ and } 130^\circ \le \phi \le 180^\circ,\ -100^\circ \le \psi \le 75^\circ\}\right]$ and right-handed $\left[\alpha_R = \{(\phi,\psi) \mid 0^\circ \le \phi \le 130^\circ,\ -180^\circ \le \psi \le 180^\circ\}\right]$ α-helical conformers. Note that other studies

split the $C7_{\mathrm{eq}}$ [144] or α_R [46] regions, or both further [29]. The classification employed here is the coarsest one still identifying the dominant conformations.

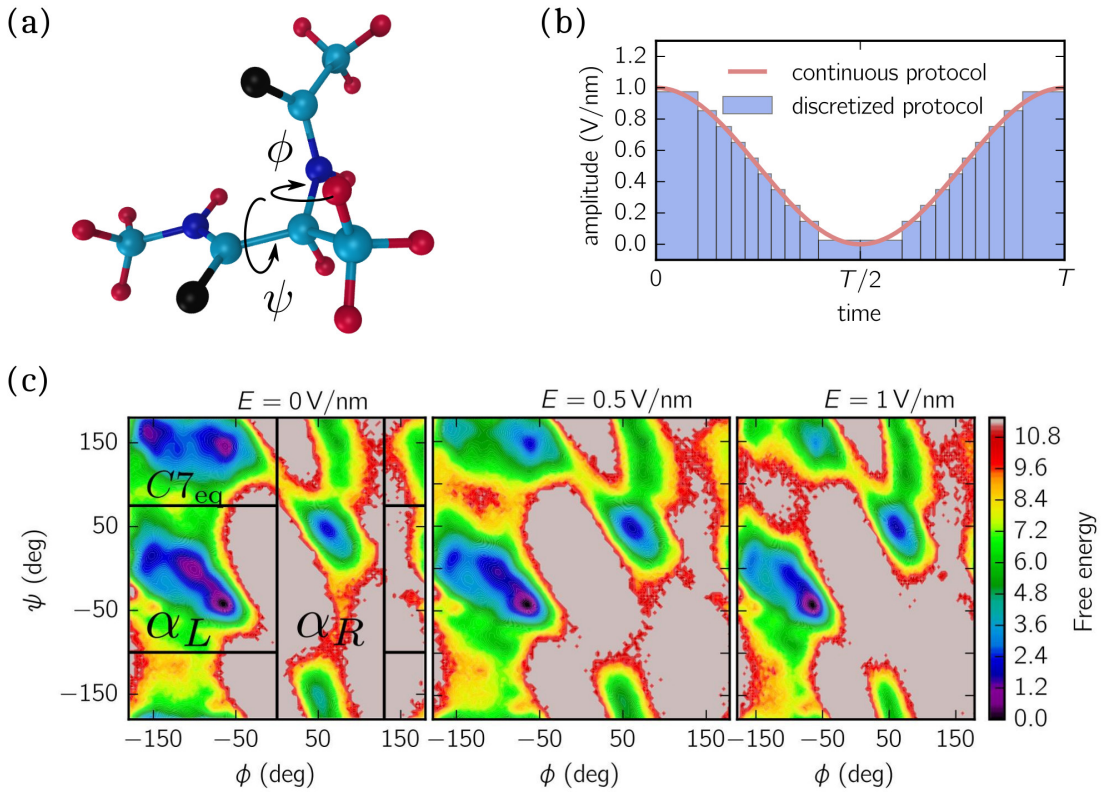

Figure 4.2 – **Alanine dipeptide.** (a) Exemplary atomistic structure of alanine dipeptide. (b) Continuous and discretized protocol $\lambda(t) = \frac{1}{2}\left[\cos\left(\frac{2\pi t}{T}\right)+1\right]$. (c) Ramachandran plot of free energies. From left to right: A constant electric field in positive x direction is applied with magnitude 1 V/nm, 0.5 V/nm and 0 V/nm. The colors in all three plots follow the same scale.

Already when applying static electric fields a significant shift in the depth of the free energy basins can be observed. For $E = 0$ V/nm the $C7_{\mathrm{eq}}$ and α_L conformations are equally dominant while the α_R conformations are comparatively rare. However, for $E = 1$ V/nm the α_R structures become more dominant while the $C7_{\mathrm{eq}}$ and α_L structures become less frequent.

4.3.3 Construction of NE-MSMs

To understand the effect of the oscillating electric field, see eq. (4.16), we construct MSMs according to the direct and indirect approach. To minimize statistical errors due to insufficient sampling, we conduct for the direct approach 1000 simulations with different starting configurations and total length of 10 ns. For the indirect approach we create 500 equilibrium trajectories for a total of 10 field strengths. To ensure that the system sufficiently equilibrated, we neglect the first 2 ns of each trajectory. In order to construct the MSM, we refer to the configuration space given by both dihedral angles (ψ, ϕ), as the dihedral angles have proven

to be good order parameters. However, instead of directly discretizing both dihedral angles, we employ the cos/sin projection for both angles, returning a four-dimensional space, i.e.,

$$\begin{pmatrix}\phi\\ \psi\end{pmatrix} \quad \longmapsto \quad \begin{pmatrix}\cos(\phi)\\ \sin(\phi)\\ \cos(\psi)\\ \sin(\psi)\end{pmatrix}. \tag{4.17}$$

The benefit of doubling the dimensionality is that, in contrast to the dihedral space, the cos/sin space offers a meaningful distance metric which allows the application of distance-based clustering algorithms.

To further improve the configuration space representation, we employ the time-lagged independent components analysis (TICA) [103].

Time-lagged independent component analysis

Similar to the principal component analysis [cf. section 2.4], TICA is an orthogonal linear transformation additionally using temporal information. We define the time lagged covariance matrices $\boldsymbol{\Sigma}(\tau)$ for which the general eigenvalue problem is solved, i.e.,

$$\boldsymbol{\Sigma}(\tau)\boldsymbol{M} = \boldsymbol{\Sigma}(0)\boldsymbol{M}\boldsymbol{D} \tag{4.18}$$

with $\Sigma_{ij}(\tau) \equiv \mathrm{cov}\left(x_i(t), x_j(t+\tau)\right)$. Here $x_i(t)$ denotes the ith component of configuration space, $\boldsymbol{M}$ the eigenvector matrix and $\boldsymbol{D}$ the diagonal eigenvalue matrix. Basically, TICA rotates the reference coordinate system so that its new axes (TICA components) point in the direction in which the slowest conformational changes occur. The projection is then obtained by only regarding the dominant TICA components which correspond to the largest eigenvalues.

For alanine dipeptide the TICA input is given by the MD trajectories projected onto the four-dimensional dihedral space [cf. eq. (4.17)]. Although the two largest largest components recover already 96% of the kinetic variance, i.e.,

$$\frac{\lambda^{(0)} + \lambda^{(1)}}{\sum_k \lambda^{(k)}} > 0.96,$$

we keep all four components as it allows the inverse transformation from TICA space back to the original two-dimensional dihedral space.

To make progress, the transformed configuration space is clustered (using the k-means algorithm) into 250 centroids [see section 1.2.1]. Centroid coordinates are labeled as $\mathbf{R}_k^{(i)}$ indicating the ith component in TICA space of the kth centroid, with $k = 1\ldots250$ and $i = 1\ldots4$.

Once all trajectories are mapped onto the centroids, we follow the direct/indirect approach to estimate the transition rate matrix $\tilde{\boldsymbol{W}}_{\mathrm{dir}}/\tilde{\boldsymbol{W}}_{\mathrm{ind}}$. For the direct approach, the transition matrix

$\boldsymbol{K}_{\mathrm{dir}}$ is first estimated for multiples of the lag time nT employing an estimator [136] that follows Bayesian inference [cf. section 1.2.2] and then converted to the final rate matrix $\tilde{\boldsymbol{W}}_{\mathrm{dir}}$.

To construct the equilibrium MSMs (for a single field strength) needed for the indirect approach, we use the maximum likelihood estimator that directly estimates $\boldsymbol{W}(\lambda_i)$ [cf. section 1.2.4]. Once all $\boldsymbol{W}(\lambda_i)$ are estimated, eq. (4.4) is solved for $\boldsymbol{K}_{\mathrm{ind}}(t=T)$ and converted to $\tilde{\boldsymbol{W}}_{\mathrm{ind}}$.

To compare both approaches, the time evolution of the probability distribution is monitored for an exemplary cumulated probability: $P_{\alpha_R}(t) = \sum_i p_i(t)$ for $(\phi,\psi) \in \alpha_R$. As shown in Figure 4.3(a), both approaches agree well with only a minor discrepancy of less than 1%.

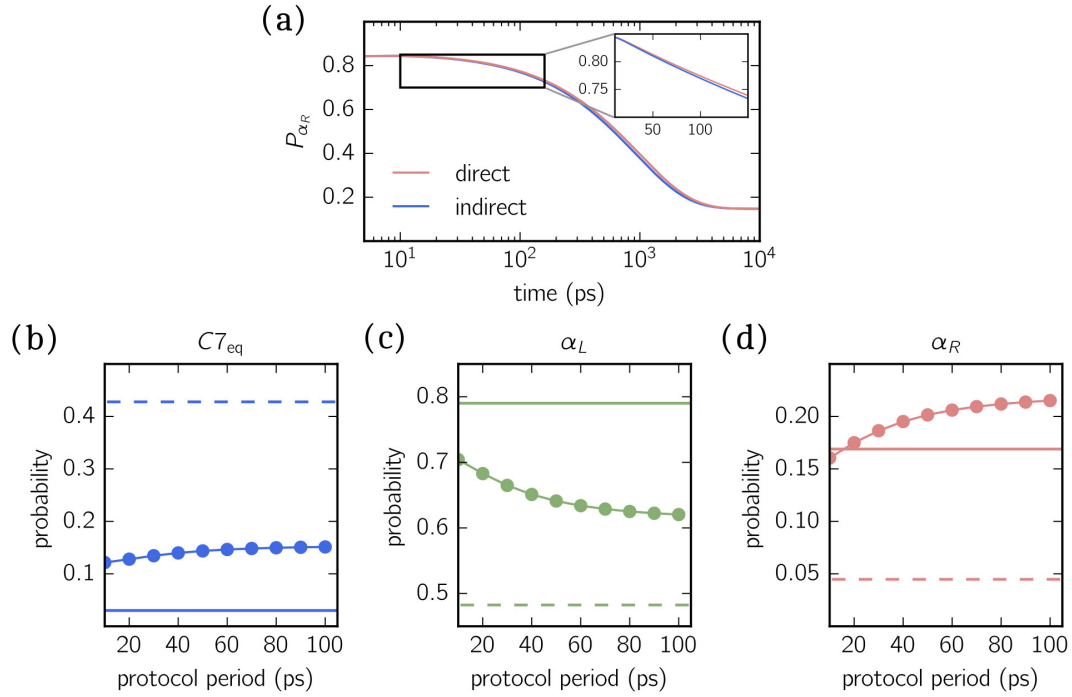

Figure 4.3 – **Time-dependent and steady-state probabilities.** (a) Comparison of the time evolution of the cumulated probability P_{α_R} following the direct and indirect approach for $T = 5$ ps. (b-d) Cumulated steady-state probabilities are shown for all three metastable sets $C7_{\mathrm{eq}}$, α_L and α_R, respectively, and different driving periods. Vertical lines correspond to equilibrium distributions for $E = 1$ V/nm (solid) and $E = 0$ V/nm (dashed).

Motivated by the good agreement of both approaches, we employ the indirect approach to predict the conformational dynamics for different driving periods. We start by investigating the influence of the period T on static properties, e.g., how the stationary probability distribution is influenced, which is interesting if one wants to suppress or enhance specific conformations. In Figure 4.3(b-d), the cumulated stationary probability distributions are shown for all metastable sets ($C7_{\mathrm{eq}}$, α_L, α_R) and different driving periods T. The solid/dashed lines illustrate the equilibrium distributions without electric field ($E = 0$ V/nm) and the strongest ($E = 1$ V/nm) statically applied electric field, cf. Figure 4.2(c). The cumulated probabilities are defined by

summing over the corresponding steady-state probabilities, i.e., $P_{C7_{\mathrm{eq}}} = \sum_i \pi_i$ for $(\phi,\psi) \in C7_{\mathrm{eq}}$ and $P_{\alpha_R/\alpha_L} = \sum_i \pi_i$ for $(\phi,\ psi) \in \alpha_R/\alpha_L$. All three distributions saturate against a fixed value for longer periods. More precisely, P_{α_L} and P_{α_R} increases while $P_{C7_{\mathrm{eq}}}$ decreases for long T. The absolute change, however, is moderate with only 10-15%. Interestingly, the behavior of the stationary distributions in comparison with the equilibrium distributions does not follow a general trend, e.g., P_{α_L} is located between both static distributions with $P_{\alpha_L}(E = 0\mathrm{V/nm}) < P_{\alpha_L} < P_{\alpha_L}(E = 1\mathrm{V/nm})$. For $P_{C7_{\mathrm{eq}}}$ the opposite trend is observed, i.e., $P_{C7_{\mathrm{eq}}}(E = 0\mathrm{V/nm}) > P_{C7_{\mathrm{eq}}} > P_{C7_{\mathrm{eq}}}(E = 1\mathrm{V/nm})$, and finally for P_{α_R} none is true as both constant values are exceeded. In particular, the latter is of importance as it demonstrates the concept of stochastic pumping [122, 12] in which an oscillating protocol is used to increase the occupation of specific states above their equilibrium values.

4.3.4 Cycle space

To gain deeper insight into the nonequilibrium dynamics, the cycle-flux decomposition [see eq. 2.8] is applied, returning all entropy producing cycles. These cycles are then represented by their cycle centers $c_\alpha^{(i)}$, diameters $d_\alpha^{(i)}$ and affinities A_α [see section 2.4], together forming the cycle space. In contrast to the previous chapters, cycle centers and diameters are computed along the two most dominant TICA components.

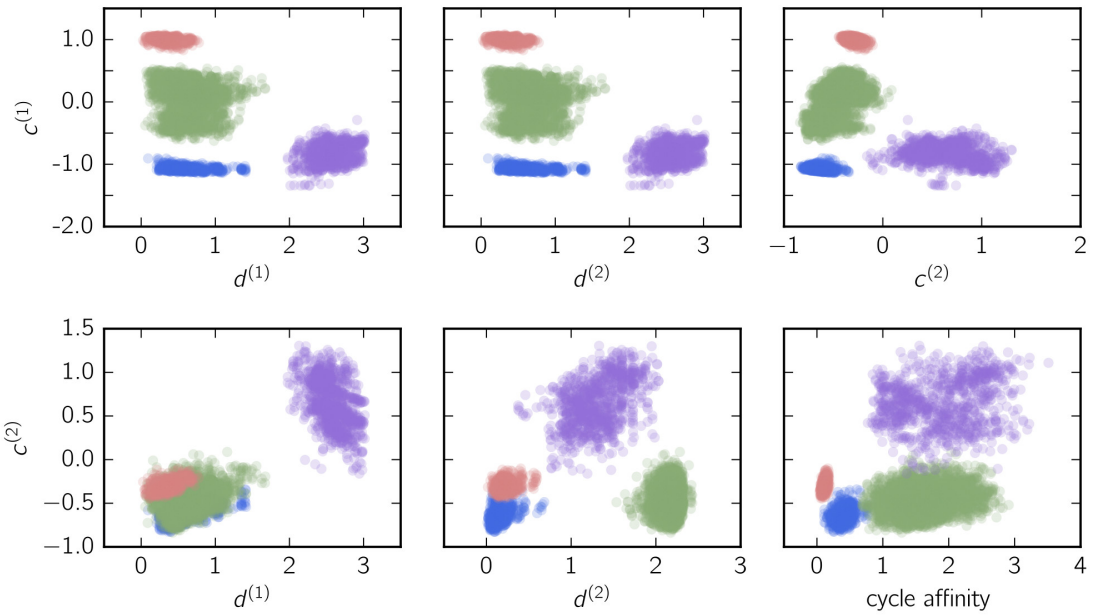

Figure 4.4 – **Cycle space projections.** Different cycle space projections are shown for $T = 5$ ps. Colors indicate detected cycle communities.

In Figure 4.4, we show six exemplary projections of these five cycle space coordinates, illustrating that the distribution of cycles exhibits an internal structure (unassigned cycles are not shown) . Our next task is to group cycles together forming cycle communities for which we

employ c-means clustering, see section 2.4. For the shown example we find $k = 5$ communities to fit the data best (highlighted by different colors).

To understand their significance concerning the change in molecular configurations, we project all cycles back onto the molecular dihedral space (ϕ, ψ) and highlight all states that are part of the same cycle community by the same color [see Figure 4.5]. Apparently, cycles and thus states belonging to the red and blue community coincide with the metastable set $C7_{\mathrm{eq}}$ and α_L, respectively. The purple and green community, on the other hand, connect each two metastable sets, with green cycles linking $C7_{\mathrm{eq}}$ and α_L configurations and purple cycles linking α_L and α_R configurations.

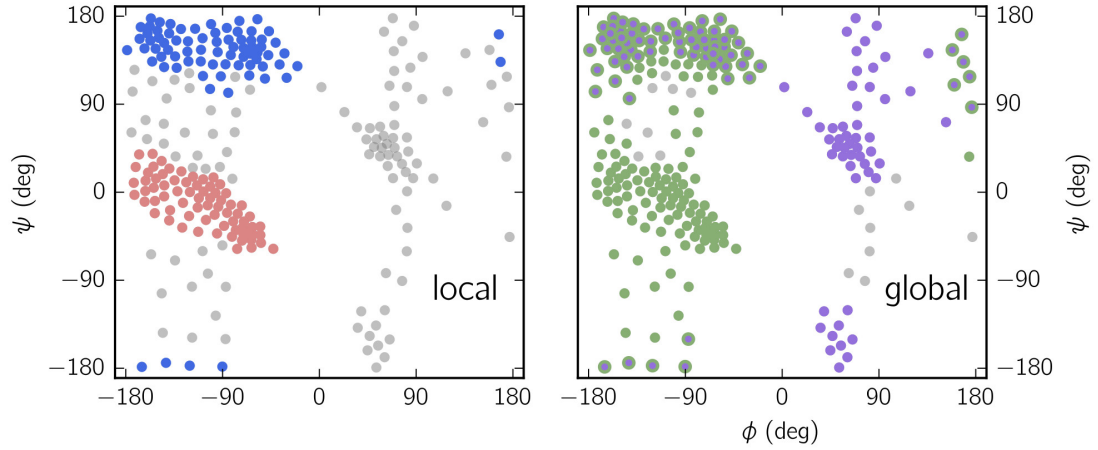

Figure 4.5 – **Cycle projection in dihedral space.** Centroids (shown in gray) of discretized TICA space are projected back onto the original dihedral space (ϕ, ψ). Colored circles illustrate centroids that belong to a specific cycle community, as shown in Figure 4.4. Centroids exhibiting two different colors (edge and face color) are part of two cycle communities. The left panel highlights all centroids belonging to local cycle communities, while the right panel highlights centroids belonging to global communities.

4.3.5 Community representatives

At this stage we understand that cycle communities occur through the interplay between the underlying potential energy landscape and the oscillating protocol that drives the system out of equilibrium. The next task is to coarse grain the NE-MSMs for which we employ the coarse graining strategy as introduced in section 2.5. To this end, all cycles forming a specific community are lumped together into a single cycle, the community representative, and all remaining cycles are deleted. The coarse graining is completed when all nonvanishing transition rates are dynamically consistently rescaled [see the rescaling algorithm in section 2.5.1] while preserving the MFPTs between all three metastable regions $(C7_{\mathrm{eq}}, \alpha_L, \alpha_R)$.

The employed coarse graining approach makes it possible to further investigate how the dynamics is altered when changing the driving period T. Therefore, we exploit the indirect

approach and construct NE-MSMs for three different periods: $T = 5$ ps, 50 ps and 100 ps. In Figure 4.6(a), the final coarse-grained rate matrices are visualized where colored cycles are representatives of their respective community. For $T = 5$ ps the coarse-grained model quantitatively confirms the cycle distribution shown in Figure 4.5. Interestingly, when the period is changed, the communities and therefore their representatives change too. For example, for $T = 50$ ps the former ($T = 5$ ps) local blue and red community vanish and two new communities, now labeled in red and blue, appear. The blue community is of importance as it encloses all three metastable sets $\alpha_R \to \alpha_L \to C7_{\text{eq}} \to \alpha_R$. Moreover, for $T = 100$ ps no new communities emerge. However, the former ($T = 5$ ps) purple and red community disappear, leaving two effective cyclic modes. The green cycle represents a clockwise rotation of the ψ angle, whereas the blue cycle connects all three metastable sets in clockwise direction. Both cycles are illustrated in Figure 4.6(b).

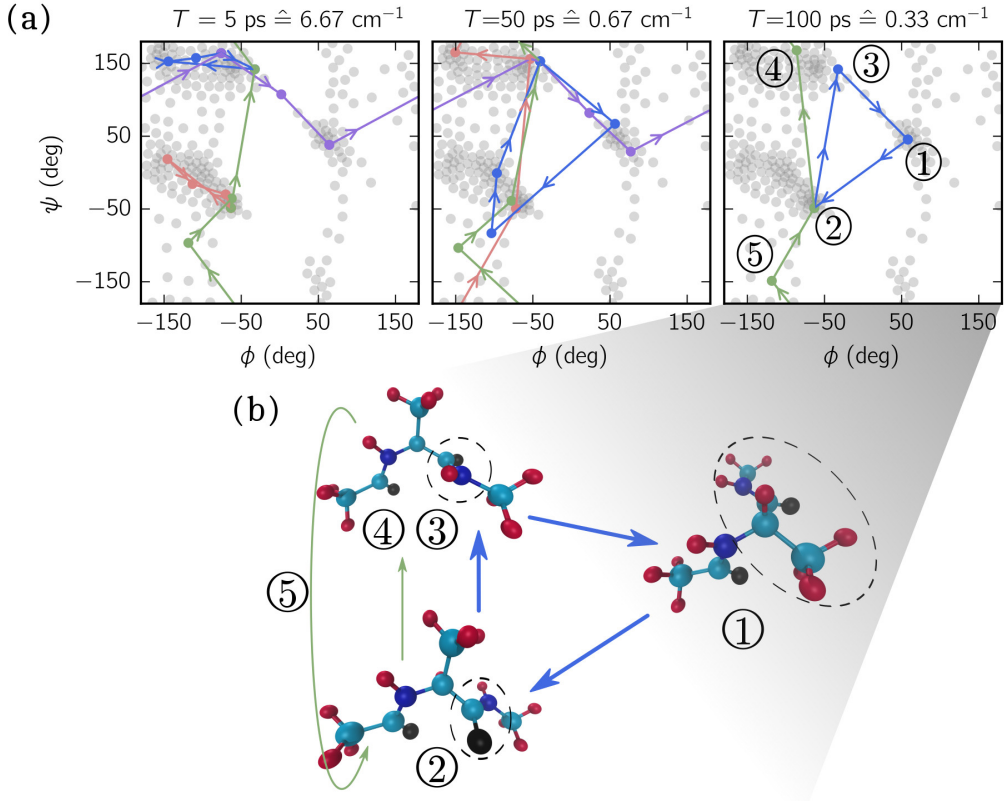

Figure 4.6 – **Cycle representatives.** (a) Final coarse-grained NE-MSM for different periods of the applied electric field. Colors indicate the community representatives. (b) Illustration of both remaining cycles for the driving period $T = 100$ ps. Dashed circles highlight the conformational changes.

Finally, we want to give an outlook about the predictive power of the indirect approach, by constructing new protocols with the same set of equilibrium MSMs that we used for discretizing the cosine protocol (see Figure 4.2b). In Figure 4.7, we show from left to right: an absolute value cosine protocol $\lambda(t) = \left|\cos(2\pi t/T)\right|$, a sawtooth protocol and a Gaussian pulse

protocol. For the last two protocols it could be interesting, besides varying the period, to vary the time between two signals.

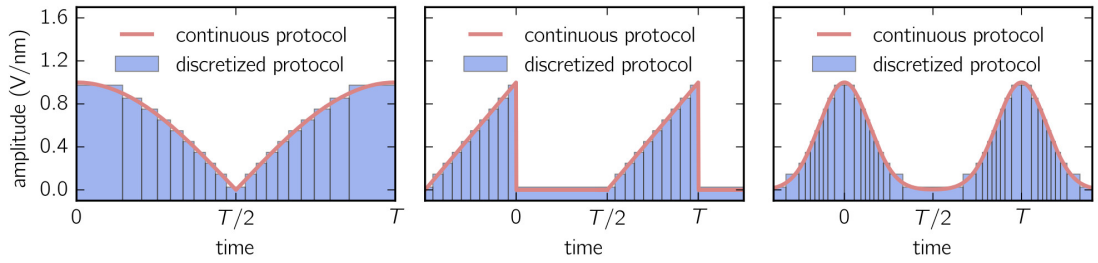

Figure 4.7 – **Periodic protocols.** All shown continuous protocols are discretized by the same set of equilibrium rate matrices, which is also used to discretize the protocol shown in Figure 4.2(b).

5 Markov State Modeling for Force Probe Simulations

Despite the fact that Markov state modeling has been proven to successfully bridge time scales from nanoseconds up to milliseconds, it still relies heavily on the molecular dynamics simulations conducted to sample the configuration space and the dynamics within. For example, if certain critical configurations are not sampled, e.g., the natively folded structure of a protein, the MSM does not contain any information about these configurations.

Even with the permanent increase in computational power, it is already a challenging task to explore and sufficiently sample the configuration space of medium-sized systems, e.g., proteins with more than 100 amino acids. To overcome this restriction, one can employ **steered molecular dynamics simulations** [55, 85, 65, 54] which accelerate the dynamics by applying biasing potentials along one or multiple reaction coordinates.

Steered MD simulations are inspired, amongst others, from the field of force probe spectroscopy where much insight into molecular conformations has been gained through atomic force microscopy [112, 27] and optical tweezer experiments [75, 28].

In these type of experiments, probing, for example, folding pathways of proteins can be realized through applying an external force to individual amino acids, which is commonly implemented by one of two different protocols: pulling or constant force (force clamp). In pulling experiments, both ends of a molecule (e.g., proteins, DNA) are separated with constant velocity or constant loading rate while recording the force-extension curve. Hence, the system is driven out of thermal equilibrium but the free energy might be recovered from the fluctuations [63, 32]. On the other hand, a constant force can be achieved through a feedback loop which, in principle, generates an equilibrium ensemble of configurations at that force. While this chapter focuses on steering molecular systems by constant-velocity protocols, constant-force simulations will be covered in the next chapter.

In steered MD simulations the constant-velocity protocol is most frequently realized by con-

necting two ends of a molecule with an additional harmonic potential

$$U_{\mathrm{bias}} = \beta \frac{\kappa}{2}\Big(R_{\mathrm{ee}}(t) - \underbrace{\left[\upsilon_{\mathrm{pull}}\ t + R^0_{\mathrm{ee}}\right]}_{R_{\mathrm{pull}}(t)}\Big)^2 = \beta \frac{\kappa}{2}\Big(R_{\mathrm{ee}}(t) - R_{\mathrm{pull}}(t)\Big)^2, \tag{5.1}$$

which resting position moves linearly in time with **pulling velocity** υ_{pull}. The biasing strength κ denotes the stiffness of the potential, $R_{\mathrm{ee}}(t)$ the distance along the bias is applied and $R^0_{\mathrm{ee}} = R_{\mathrm{ee}}(0)$ this distance at $t = 0$. We further define $R_{\mathrm{pull}}(t) \equiv \upsilon_{\mathrm{pull}}\ t + R^0_{\mathrm{ee}}$ as the absolute resting position. Although the biasing potential accelerates the dynamics, for moderate system sizes the accessible pulling velocities in computer simulations are of the order $\upsilon_{\mathrm{pull}} \sim 1$ m/s, which is more than 10^6 times faster than pulling velocities typically realized in experiments. This large discrepancy makes it difficult to compare simulations and experiments, in particular as, for example, the unfolding pathways of proteins can depend on the pulling velocity [150, 111].

In this chapter, we introduce the concept of Markov state modeling to steered molecular dynamics simulations, bridging the gap between numerically and experimentally reachable pulling velocities. We illustrate our methodology for the Calix[4]arene-catenane dimer, a large organic complex.

5.1 Constant-velocity pulling

Incorporating an explicitly time-dependent biasing potential [see eq. (5.1)] causes the rate matrix to be time-dependent. Due to this time dependency, the master equation

$$\frac{\partial p^{\mathrm{T}}(t)}{\partial t} = p^{\mathrm{T}}(t)\ \boldsymbol{W}(t) \tag{5.2}$$

does not have a stationary solution. However, similar to the indirect approach of the previous chapter, the biasing potential can be discretized $U_{\mathrm{bias}}(\lambda(t))$ by discretizing the **pulling protocol** $\lambda(t) \equiv R_{\mathrm{pull}}(t)$, which allows the construction of equilibrium MSMs for given $\lambda(t = t_i) = \lambda_i$. Here the protocol λ has the physical meaning of a pulling distance.

When the protocol λ is frozen in time, the biasing potential $U_{\mathrm{bias}}(t = t_i)$ together with the regular potential energy U form an effective time-independent potential. Therefore, for a fixed point in time the rate matrix $\boldsymbol{W}(\lambda_i)$ can be estimated by employing the standard equilibrium Markov state modeling machinery [cf. section 1.2]. Knowing the rate matrices $\boldsymbol{W}(\lambda_i)$ allows the master equation to be solved until final time t_F. Most importantly, the pulling velocity υ_{pull} only enters in the last step. Thus, changing υ_{pull} only requires eq. (5.2) to be solved again.

Before applying this procedure to a model system, we introduce the **transition-based reweighting method**.

5.2 The transition-based reweighting method

The transition-based reweighting method (TRAM) is an approach to connect multiple equilibrium ensembles, moving beyond well-established approaches such as the weighted histogram analysis method [79]. Its main advantage is that rare events in one ensemble can become less rare in another ensemble. TRAM exploits information about the probability distribution sampled in these "less rare" ensembles to improve estimations of rare events. Here we only give a brief summary of the method and refer for a detailed description to the original publications [90, 149].

We consider a system effectively described by an one-dimensional order parameter r (the generalization to multi-dimensional order parameters is straightforward). Defining the dimensionless potential of mean force as $\beta\, U(r)$, the probability distribution is given by the Boltzmann factor

$$\mu(r) = \frac{e^{-\beta\, U(r)}}{Z}, \tag{5.3}$$

with Z denoting the partition function. If a biasing potential is applied along the order parameter, the potential of mean force becomes $U^k(r) = \beta\, U(r) + B^k(r)$ with index k denoting the kth thermodynamic ensemble, and $B^k(r)$ any differentiable function.

Here we define the biasing potential as

$$B^k(r) = \frac{\beta\,\kappa}{2}\left(r - R^k_{\text{pull}}\right)^2, \tag{5.4}$$

where r denotes the end-to-end distance R_{ee} and R^k_{pull} the absolute pulling distance along the direction of R_{ee}. Both distances are illustrated for our model system in Figure 5.1(a).

The probability distribution of the kth ensemble is accordingly expressed as

$$\mu^k(r) = p(r)\,\frac{e^{-B^k(r)}}{Z^k}, \tag{5.5}$$

where $Z^k = Z\big(R^k_{\text{pull}}\big)$ is the partition function of the kth biased ensemble.

For every ensemble one can then construct equilibrium MSMs (obeying detail balance), for which the configuration space is discretized and all observed transitions are counted, e.g., c^k_{ij}. For the purpose of readability, we switch temporarily the index notation, i.e., indices denoting transitions are now given by subscripts, e.g., $T_{ij} = T^i_j$.

When ignoring the information provided by the other ensembles, the transition probabilities T^k_{ij} can be estimated by employing the maximum likelihood estimator

$$L^k = \prod_{i,j}\left(T^k_{ij}\right)^{c^k_{ij}}, \tag{5.6}$$

as introduced in section 1.2.2.

To infer information from multiple ensembles, a local equilibrium condition for the stationary probability of states (volume cells) S_i is defined as

$$\mu_i^k(r) = \begin{cases} \mu^k(r)\ \frac{Z^k}{Z_i^k} & r \in S_i \\ 0 & \text{else} \end{cases}, \tag{5.7}$$

with

$$Z_i^k = Z^k \int_{S_i} \mu^k(r)\ \mathrm{d}r \tag{5.8}$$

being the local partition function belonging to states S_i. Eq. (5.7) holds for the milder and thus advantageous assumption that the underlying MD simulations do not need to be sampled from a global equilibrium distribution covering all S_i's.

Connecting eq. (5.5) with eq. (5.6) and exploiting the local equilibrium condition [eq. (5.7)], the TRAM maximum likelihood function follows as

$$L_{\mathrm{TRAM}} = \prod_k \left(\prod_{i,j} \left(T_{ij}^k \right)^{c_{ij}^k} \right) \left(\prod_i \prod_{r \in S_i^k} \mu(r)\ \frac{e^{-B^k(r)}}{Z_i^k} \right), \tag{5.9}$$

where S_i^k denotes the set of all samples drawn from the kth ensemble and assigned to state S_i. Eq. (5.9) can be further simplified and efficiently numerically maximized. Here we use an implementation provided in the PyEMMA software package [117].

5.3 Model system

As a model system we choose the calix[4]arene-catenane dimer that is composed of two cuplike structures formed by four benzene molecules which are connected via alkyl groups in a ringlike fashion [19]. A snapshot visualizing the atomistic structure is depicted in Figure 5.1(a), where two groups of atoms are highlighted in red and blue which define the reference and pulling group, respectively. In more detail, both groups include four carbon atoms of the methoxy groups at the narrow rim of one of the calix[4]arene molecules. The distance connecting the center of mass of each group is referred to as end-to-end distance R_{ee}, which we adopt as an order parameter. The spring illustrated in Figure 5.1(a) represents the biasing potential defined in eq. (5.1). The pulling distance R_{pull} denotes the absolute distance between the reference group and the spring or rather its resting position. Note that for $R_{\mathrm{pull}} = R_{\mathrm{ee}}$ the biasing potential is zero and thus no biasing force is applied.

An exemplary trajectory illustrating the pulling protocol for $v_{\mathrm{pull}} = 1$ m/s is shown in Figure 5.1(b). At the beginning, the end-to-end distance follows linearly the pulling protocol, although exhibiting a much smaller slope. At $t \approx 2.5$ ns, the end-to-end distance jumps to

a larger value, indicating that the cuplike structures of the calix[4]arene dimer are largely separated. After this sharp transition, R_{ee} grows again linearly with the pulling protocol. The calix[4]arene dimer is referred to as being in the closed state when both cuplike structures are close together, and being in the opened state if they are largely separated. The abrupt jump in R_{ee} marks the transition between both states.

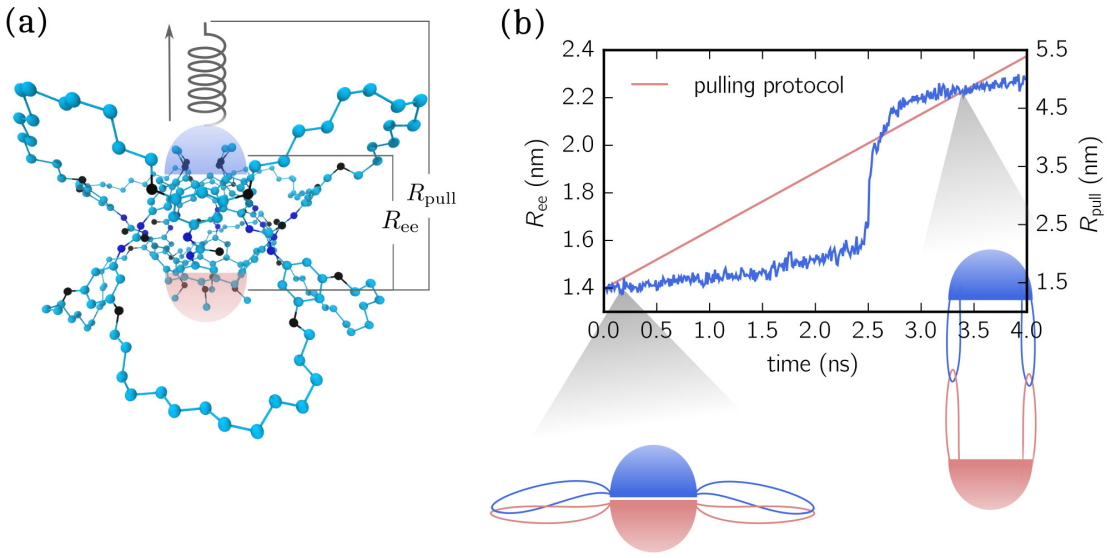

Figure 5.1 – **Calix[4]arene-catenane dimer.** (a) Atomic structure of calix[4]arene. The reference and pulling group are highlighted by a red and blue cup, respectively. During the simulation, the pulling group is separated from the reference group by a moving harmonic potential. (b) End-to-end distance and pulling protocol for $v_{pull} = 1$ m/s. Additionally, sketches illustrating the opened and closed conformations of the calix[4]arene dimer are shown.

Schlesier *et al.* [119, 118] found the sharp transition of the end-to-end distance to be consistent with the rupture and formation of hydrogen bonds between urea groups and between urea and ether oxygen groups, respectively. Moreover, the authors conducted many independent simulations showing that the jumping time (when the molecule switches from the closed state to the opened state) is indeed a stochastic event.

5.3.1 Modeling details

MD simulations

All molecular dynamics simulations are performed employing the Gromacs 5.1.2 [1] software package. The calix[4]arene dimer is positioned in a 5.7 nm×5.7 nm×5.7 nm box with periodic boundary conditions and filled with 827 mesitylene molecules. Molecular interactions are described by the OPLS [72] force field. All simulations are conducted in the NPT ensemble, for which the temperature is set to 300 K employing the velocity-rescaling thermostate [25]. To maintain a constant pressure of one bar, the Parrinello-Rahman [102] barostat with $\tau_p = 2$ ps and compressibility of 8.26×10^{-5} bar^{-1} is used. Long-range electrostatics are treated using

particle mesh Ewald summation method [36], while for van der Waals interactions a dispersion correction [5] is applied. The cutoff for all short-ranged interactions is set to 1.4 nm. All hydrogen involving covalent bonds are constraint by the LINCS algorithm [58], allowing a time step of 2 fs.

Overall, we perform multiple pulling simulations including three different pulling velocities $v_{\text{pull}} = \{0.01\ \text{m/s}, 0.1\ \text{m/s}, 1\ \text{m/s}\}$. For the construction of MSMs, we approximate R_{pull} 41 values (differing by 0.05 nm) placed between 2.0 nm $\leq R_{\text{pull}} \leq$ 4.0 nm. Between 40-100 MD simulations with an individual length of 20 ns are conducted for fixed (time-independent) biasing potential $U_{\text{bias}}\left(R^k_{\text{pull}}\right)$ [see eq. (5.1)]. Throughout all simulations, the stiffness of the biasing potential is set to $\kappa = 500\ \text{kJ/(mol·nm}^2)$.

All MD simulations have been conducted by Ken Schäfer.

Markov state modeling

We start by discretizing the one-dimensional order parameter R_{ee} into 60 centroids (states) with respective positions R^i_{ee} equally placed between 1.35 nm and 2.2 nm. Next, the trajectories for all 41 biasing potentials $U_{\text{bias}}\left(R^k_{\text{pull}}\right)$ are mapped onto the set of discrete states and the number of observed transitions is counted for each thermodynamic ensemble.

Employing TRAM [cf. section 5.2] with lag time $\tau_{\text{TRAM}} = 1$ ns, a set of estimated transition matrices $\boldsymbol{T}^k$ is returned and converted, following eq. (1.51), to rate matrices $\boldsymbol{W}^k$.

5.3.2 Biased free energy landscapes and mean first passage times

Before reconstructing the pulling protocol, we examine the static and dynamical properties of the time-independent rate matrices $\boldsymbol{W}^k$.

In Figure 5.2(a), selected free energies along the end-to-end distance are depicted for different values of R_{pull}. The free energy landscapes follow from the stationary distributions π^k via $F^k_i = -k_B T\ \log\left(\pi^k_i\right)$ and are rescaled by setting their respective value for $R_{\text{ee}} = 1.48$ nm to zero. For $R_{\text{pull}} = 2.8$ nm the global minimum coincides with the closed state [$R_{\text{ee}} < 1.5$ nm], while the remaining part of the free energy is rather flat [$R_{\text{ee}} > 1.7$ nm]. For $R_{\text{pull}} = 3.1$ nm, on the other hand, the global minimum is shifted toward the opened state [$R_{\text{ee}} > 1.8$ nm] with the closed state still exhibiting a distinct local minimum. Intermediate values of R_{pull} illustrate the transition between both cases. In particular, for R_{pull} between 2.95 nm and 3.0 nm the opened and closed state belong to equally deep minima, which are separated by the lowest free energy barrier [at $R_{\text{ee}} \approx 1.7$ nm] found for all biasing potentials.

Knowing the rate matrices $\boldsymbol{W}^k$ allows us to compute mean first passage times connecting the closed and opened state. In more detail, all states with centroid positions $R_{\text{ee}} < 1.5$ nm constitute the closed state, whereas states with $R_{\text{ee}} > 1.8$ nm constitute the opened state.

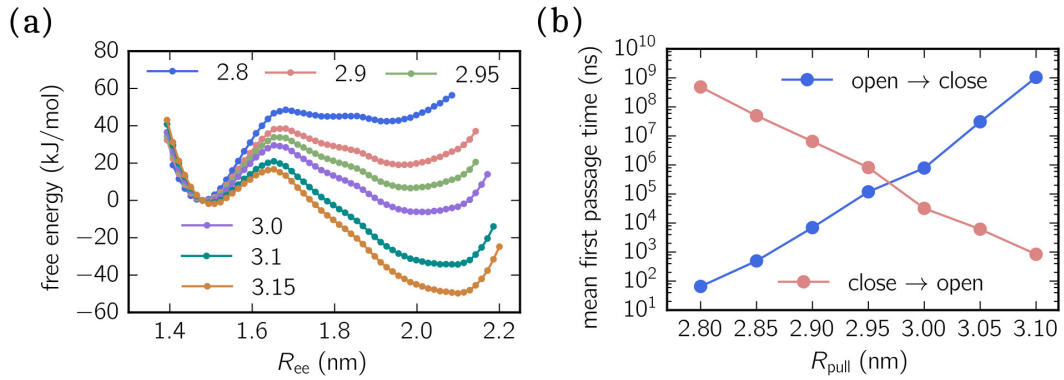

Figure 5.2 – **Free energy landscapes and mean first passage times.** (a) The free energy landscape along the end-to-end distance R_{ee} is shown for different positions R_{pull} (in nm) of the biasing potential. (b) Mean first passage times between opened ($R_{ee} > 1.8$ nm) and closed ($R_{ee} < 1.5$ nm) configurations for different positions R_{pull}.

The MFPTs between both sets of states are shown in Figure 5.2(b) for different positions R_{pull}. When plotting in a semilogarithmic scale, the MFPTs follow a line and therefore are in agreement with Kramer's rate theory [see section 1.2.5], for which the free energy barrier that separates both states dictates the transition time scale. Moreover, both MFPTs cross at $R_{pull} = 2.975$ nm, which is in excellent agreement with the particular free energy landscape exhibiting equally deep free energy minima.

One should note that both MFPTs cover time scales from tenth of nanoseconds up to tenth or even hundreds of milliseconds, which is a remarkable result of the TRAM estimator, especially when taking into account that every MSM is constructed from a total of 2 μs or less of accumulated trajectories.

5.3.3 Reconstructing the pulling protocol

Next, we reconstruct the pulling protocol $\lambda(t)$ by approximating the time-dependent rate matrix $\boldsymbol{W}(t)$ with the set of time-independent rate matrices $\boldsymbol{W}(\lambda_k)$, yielding an approximated solution $p_{\mathrm{MSM}}(t)$ of the master equation, i.e.,

$$p_{\mathrm{MSM}}^{\mathrm{T}}(t_F) = p_0^{\mathrm{T}} + \sum_{k=0}^{N} \int_{t_k}^{t_{k+1}} \mathrm{d}t \; p^{\mathrm{T}}(t) \; \boldsymbol{W}(\lambda_k). \tag{5.10}$$

The pulling velocity v_{pull} is then determined by

$$v_{\mathrm{pull}} = \dot{\lambda} = \frac{\lambda_{N+1} - \lambda_0}{t_F}, \tag{5.11}$$

where λ_k represents the kth pulling distance R_{pull}^k. As we cannot compare $p_{\mathrm{MSM}}(t)$ with the true solution, we compare the average jumping time obtained from pulling simulations [cf. Figure 5.1(b)] with the jumping time following the reconstructed pulling protocol. In

addition, the jumping time serves as a dynamical quantity which is in principle experimentally accessible [66]. To make progress, we compute the time-dependent mean end-to-end distance via

$$\langle R_{\mathrm{ee}}^{\mathrm{MSM}}(t)\rangle = \sum_i R_{\mathrm{ee}}^i \, p_{\mathrm{MSM}}^i(t) \quad \text{and} \quad \langle R_{\mathrm{ee}}^{\mathrm{MD}}(t)\rangle = \sum_j R_{\mathrm{ee}}^j(t) \tag{5.12}$$

for the Markov state modeling approach and for the pulling MD simulations, respectively. Here $\langle R_{\mathrm{ee}}^{\mathrm{MSM}}(t)\rangle$ is determined by computing the ensemble average, where R_{ee}^i denotes the ith centroid position [see section 5.3.1] and $p_{\mathrm{MSM}}^i(t)$ its associated probability at time t. For the pulling simulations, $\langle R_{\mathrm{ee}}^{\mathrm{MD}}(t)\rangle$ is determined by averaging over all conducted MD simulations.

Both time-dependent mean end-to-end distances are shown for $v_{\mathrm{pull}} = 0.01$ m/s in Figure 5.3(a). Although $\langle R_{\mathrm{ee}}^{\mathrm{MD}}(t)\rangle$ exhibits small irregularities (we averaged over 22 MD trajectories), its general behavior agrees well with $\langle R_{\mathrm{ee}}^{\mathrm{MSM}}(t)\rangle$, i.e., both show a continuous transition between the closed and opened state. To quantify the jumping time τ^*, both curves are fitted by

$$f(t) = \frac{a_0}{2}\left(\tanh\left(a_1\,[t-\tau^*]\right)+1\right) + b_0\,t + b_1. \tag{5.13}$$

The inverse jumping time $1/\tau^*$ then yields the unfolding rate.

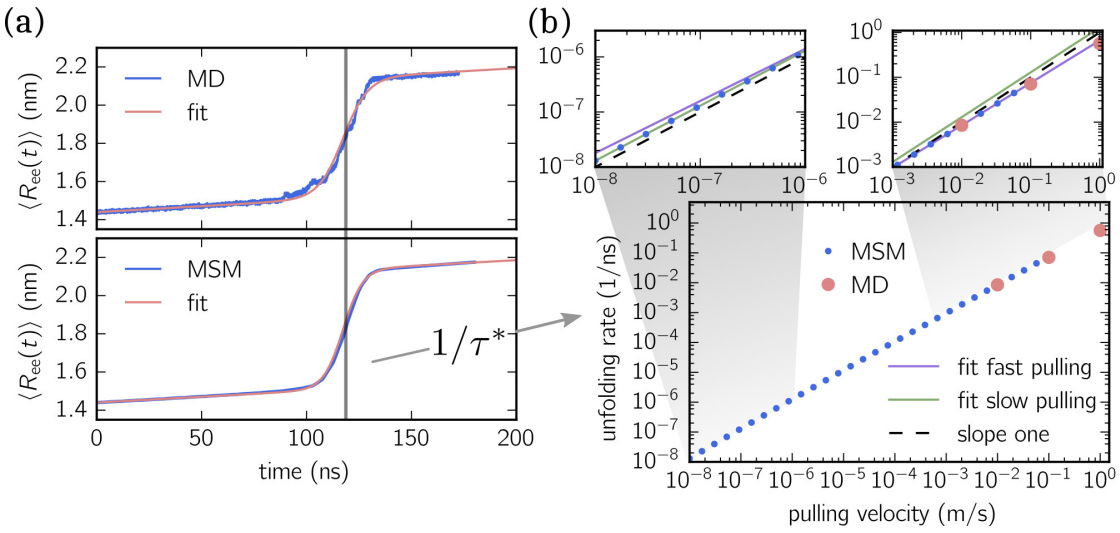

Figure 5.3 – **Mean end-to-end distance and unfolding rates.** (a) The time-dependent mean end-to-end distance is shown for $v_{\mathrm{pull}} = 0.01$ m/s, computed from 22 MD simulations (upper panel) and computed from the NE-MSM (lower panel). The red curves follow a nonlinear fit [see eq. (5.13)]. (b) Unfolding rate for different pulling velocities. A detailed description how the unfolding rates are obtained is given in the main text. Upper panels show the detailed unfolding rates for slow (left panel) and fast (right panel) pulling velocities. Solid lines represent fits for $v_{\mathrm{pull}} > 10^{-3}$ m/s (purple) and $v_{\mathrm{pull}} < 10^{-6}$ m/s (green), while the black dotted line has slope one.

In Figure 5.3(b), we compare the unfolding rates for different pulling velocities v_{pull} for both approaches. The unfolding rates for $v_{\text{pull}} = \{0.01, 0.1\}$ m/s following the discretized pulling protocol are in perfect agreement with the unfolding rates obtained from steered MD simulations. Note that the fitting error is smaller than the marker. Unfortunately, we can only compare unfolding rates for 0.01 m/s$\leq v_{\text{pull}} \leq$ 0.1 m/s as slower pulling velocities become inaccessible for brute-force MD simulations. Faster pulling velocities, i.e.,$v_{\text{pull}} >$ 0.1 m/s, on the other side, become inaccessible for the Markov state modeling approach, which limitations we will discuss at the end of this chapter.

Employing the discretized protocol, we determine the unfolding rates for pulling velocities as low as $v_{\text{pull}} = 10^{-8}$ m/s. Albeit all unfolding rates plotted in log-log scale appear to follow a straight line, a closer look suggests that for slower pulling the curve exhibits a different slope (exponent) than for faster pulling. To quantify this difference, we fit the unfolding rates by $g(x) = g_0\ x^a$ for $v_{\text{pull}} < 10^{-6}$ m/s and $v_{\text{pull}} > 10^{-3}$ m/s, respectively, which is illustrated in the insets in Figure 5.3(b). The fitted exponents are

$$a_{\text{slow}} = 0.999 \quad \text{and} \quad a_{\text{fast}} = 0.945, \tag{5.14}$$

stating that for slow pulling the unfolding rates match the pulling velocities.

5.3.4 Limitations

Theoretically, the protocol $\lambda(t)$ can be approximated for all pulling velocities. In practice, however, for certain pulling velocities the discretization is not fine enough or the Markov assumption is violated.

The **upper limit** for the pulling velocity is determined by two factors. First, for large values of R_{pull} the sampled configuration space does not cover the full centroids space, for example $\boldsymbol{W}(R_{\text{pull}} = 2.7\text{ nm}) \in \mathbb{R}^{60\times60}$ but $\boldsymbol{W}(R_{\text{pull}} = 3.6\text{ nm}) \in \mathbb{R}^{52\times52}$. For the latter, transition rates linking centroids with lower values of R_{ee}, i.e., $R_{\text{ee}} < 1.47$ nm are missing. However, for moderate and slow pulling velocities the error is negligible. When pulling too fast, on the other hand, probability is trapped at small end-to-end distances. Second, one should note that the set of rate matrices $\boldsymbol{W}^k$ is estimated for a certain lag time. If the pulling velocity and thus the final time t_F (for which the master equation is solved) is shorter than the lag time, the Markov assumption is not guaranteed to hold anymore.

The **lower limit**, on the contrary, is dictated by the slowest time scale of the set of rate matrices. When the pulling velocity is lower than the slowest time scale, the reconstructed dynamics is quasistatic and thus perfectly following the pulling protocol. However, for the true dynamics this does not need to be the case. To determine the lower limit for a given discretization of R_{pull} and set of recorded trajectories, we compute the intersection between both (slow and fast pulling) fits [see insets in Figure 5.3(b)], for which we find $v^*_{\text{pull}} \approx 4 \times 10^{-6}$.

To conclude, our approach is able to determine the unfolding rate for pulling velocities many

orders of magnitude lower than accessible by brute-force MD simulations, which allows a comparison with experiments.

6 Unfolding dynamics of small peptides biased by constant mechanical forces

In this chapter, we focus on the second type of force probe simulation: constant-force simulations [100, 128, 17]. Instead of applying a time-dependent biasing potential, in constant-force simulations the dynamics is accelerated through a time independent constant biasing force. Its advantage is that all molecular interactions can be described by an effective Hamiltonian. Therefore, constant-force simulations follow equilibrium statistical mechanics, allowing the construction of equilibrium MSMs.

The biasing potential separating two ends of a molecule is expressed as

$$U_{\text{bias}} = -\beta\, f\, R_{\text{ee}}, \tag{6.1}$$

where R_{ee} denotes the end-to-end distance connecting between both ends. The factor f represents the strength of the biasing potential and will indicate different constant-force ensembles.

Here we illustrate the application of equilibrium Markov state modeling to constant-force simulations for two specific peptides.

6.1 Molecular dynamics simulation

For all MD simulations we employ the Gromacs 5.1.2 software package [1]. All simulations are conducted in the NPT ensemble with temperature set to 300 K, for which the velocity-rescaling [25] thermostat with $\tau_T = 1$ ps is employed. For isotropic pressure coupling, we use the Parrinello-Rahman [102] barostat with $\tau_p = 2$ ps and compressibility of 4.5×10^{-5} bar^{-1}. Long-range electrostatics are treated employing particle mesh Ewald summation [36] with cubic interpolation and Fourier grid spacing of 0.16 nm. All short-ranged interactions are cutoff at 1.0 nm and all hydrogen-involving covalent bonds are restrained by the LINCS algorithm [58]. The time step is set to 2 fs.

6.2 Example I: Deca-alanine

As a first example we choose the well studied peptide deca-alanine, which is an alpha-helical forming model systems [57, 69]. To study how the folding/unfolding dynamics changes, a constant-force is imposed between the center of mass of the N-terminus and the center of mass of the C-terminus of the peptide, the distance between which defines the end-to-end distance of the peptide.

We create 48 independent simulations from multiple starting configurations for four different force magnitudes $f = \{0, 5, 10, 15\}$ kJ/mol·nm, see eq. (6.1), accumulating a total of $2-2.5$ μs per force. The molecular interactions of deca-alanine are described by the CHARMM22/CMAP [86] force field, while we employ TIP3P [71] water to model solvent interactions. The peptide is positioned in a cubic box with 7.2 nm box length and periodic boundary conditions, solvated by $\simeq$ 12000 water molecules. To ensure that every trajectory is properly equilibrated we discard its first 2 ns.

To construct a reasonable set of order parameters, we employ the TICA algorithm [see section 4.3.3]. As original configuration space representation serving as input for TICA, we choose all C_α pair distances which are separated by at least two residues and, additionally, all hydrogen bond distances between donor (nitrogen) and acceptor (oxygen) atoms that stabilize alpha-helical structures. More precisely, we use the distance between C=O of the ith residue and N–H of the $(i+4)$th residue. For this 34-dimensional configuration space, we keep the first four TICA components exhibiting a cumulative kinetic variance $> 92\%$. Together with the end-to-end distance, the 5-dimensional reduced configuration space is discretized employing the k-means clustering algorithm with 300 clusters.

After mapping the trajectories of all constant-force ensembles onto the reduced configuration space and counting all observed transitions, we use TRAM [see section 5.2] for a lag time of 4 ns returning a fine-grained MSM for every ensemble containing 300 micro-states (clusters).

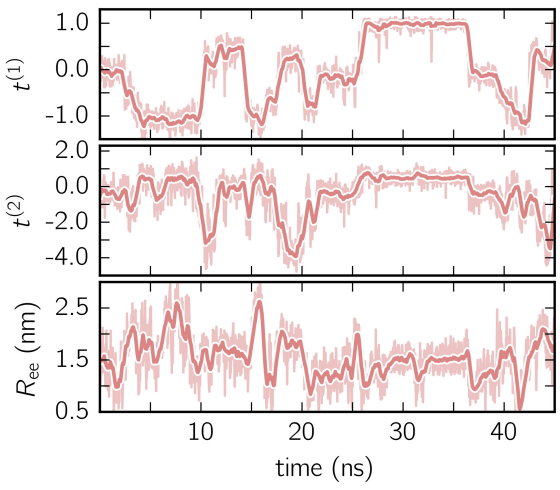

Figure 6.1 – **Exemplary times series of deca-alanine.** All three panels show the same time series projected onto the first TICA component (upper panel), the second TICA component (center panel) and end-to-end distance (lower panel). Detailed time series are illustrated by shaded lines, whereas thick lines represent moving averages.

In Figure 6.1, an exemplary time series of the first two TICA components and the end-to-end distance for deca-alanine is shown. First, note that more than two states are involved which can be clearly seen in the time series of the first TICA component (upper panel) illustrated by jumps between plateaus (metastable states) with different amplitudes. The time series of the second TICA component offers less plateaus with more fluctuations (see the shaded time series), while for the time series of the end-to-end distance (lower panel) plateaus are insufficiently resolved. Additionally, fluctuations of the end-to-end distance are much stronger as for the TICA components shown. The comparison of all three time series illustrates that for deca-alanine both TICA components are better suited as order parameters than the end-to-end distance of the peptide.

6.2.1 Configuration space

To gain insight into stable configurations of deca-alanine, we make use of the PCCA+ algorithm, as introduced in section 1.3. In more detail, we employ the PCCA+ algorithm for the transition matrix belonging to the unbiased ensemble and keep the assignment of metastable sets throughout the other ensembles. The transition matrix shows a spectral gap between its 6th and 7th largest eigenvalue (not shown) implying a time scales separation. To exploit this time scale separation, we choose $k = 6$ as input parameter for PCCA+ algorithm.

In Figure 6.2, two projections of the discretized configuration space are shown along the first two TICA components as well as along the first TICA component and the end-to-end distance. While different colors indicate different metastable sets, gray centroids belong to transition states. One issue arises here: Although the PCCA+ suggests 6 metastable sets, we manually separate stretched configurations (light blue colored set) from unfolded/weakly coiled configurations (brown set) as the former can be clearly identified by their end-to-end distance. PCCA+ does not detect stretched configurations automatically because they are temporally not well separated from the remaining unfolded configurations (brown set). These remaining unfolded configurations are weakly coiled along the centered residues, while outer residues can fluctuate freely. Large fluctuations of the outer residues can lead to hairpin-like structures which are stabilized by hydrogen bonds of the outer residues (yellow set). Hairpin-like structures are misfolded configurations acting as traps which slow down the folding process. Other misfolded configurations are detected by the red, purple, blue and cyan colored sets. All of these misfolded configurations have in common that at some residue the helical structure is disturbed, e.g., coiled in the wrong direction and stabilized by hydrogen bonds different from the ones guaranteeing the correct alpha-helical structure. Lastly, the green metastable set represents correctly folded configurations.

6.2.2 Free energy landscape and unfolding/folding rates

To gain quantitative insight and to compare different constant-force ensembles, we compute the free energy landscape along the end-to-end distance, which is obtained by computing

the stationary solution of the MSMs, i.e., $F_i = -k_B T \log \pi_i$. Figure 6.3(a) (upper panel) shows the free energy of every micro state with ($f = 10$ kJ/mol·nm) and without constant-force bias ($f = 0$ kJ/mol·nm). The global minimum for both ensembles is found at $f_{\min} \approx 1.6$ nm, which corresponds to the alpha-helical structure.

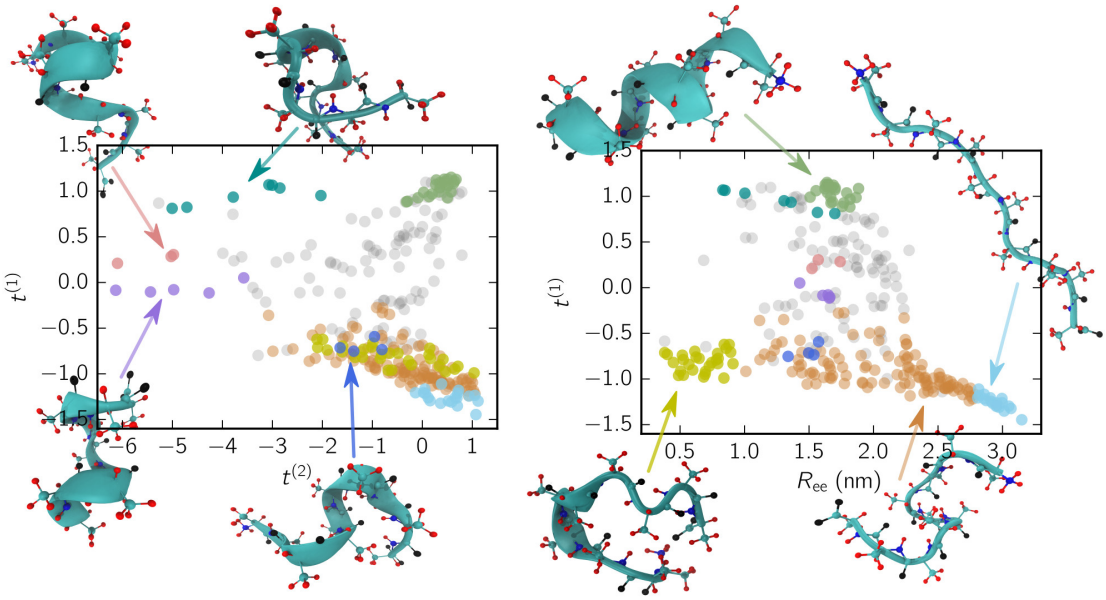

Figure 6.2 – **Configuration space of deca-alanine.** Configuration space projected onto the first and second TICA component (left panel) and onto the first TICA component and the end-to-end distance (right panel). Circles correspond to centroids discretizing configuration space (fine graining). Colored circles represent configurations belonging to metastable sets with different colors corresponding to different metastable sets. An exemplary snapshot for every metastable configuration is shown and related to its position in configuration space. Gray circles represent transition states which are not assigned to any metastable set.

Although the position of the minimum is identical for both ensembles, their free energy landscapes differ drastically. While for $f = 0$ kJ/mol·nm the free energy for shorter end-to-end distances is lower than for elongated configurations, the opposite applies for $f = 10$ kJ/mol·nm where the free energy of elongated configurations is almost as low as for folded ones. Furthermore, the maximal free energy difference is up to 25% higher for the unbiased ensemble. To have a better comparison, we integrate out all dimensions except the end-to-end distance. The results for all ensembles are shown in Figure 6.3(b) (lower panel) for which the probability distributions are binned with a bin size of 0.15 nm. For $f = \{0, 5, 10\}$ kJ/mol·nm the global minimum is located at the same position as in Figure. 6.3(a). However, for $f = 15$ kJ/mol·nm the global minimum is shifted ($\simeq 2.5$ nm) toward the elongated configurations.

Knowing the probability distribution for all ensembles, we determine the cumulated probability of populating misfolded states [cyan, yellow, blue and purple colored states in Figure 6.2]. As show in Figure 6.4(a), all probabilities decrease drastically with increasing biasing force

f. The sum over all misfolded probabilities (inset) equals approximately 17% without bias and vanishes (less than 1%) for $f = 15$ kJ/mol·nm. Individually, cyan colored configurations exhibit the largest population probability for $f = 0$ kJ/mol·nm (approx. 8%), followed by yellow and purple colored configurations, with red and blue colored states being less important. For $f = \{5, 10\}$ kJ/mol·nm the order is partially disturbed and lost (or not significant) for $f = 15$ kJ/mol·nm.

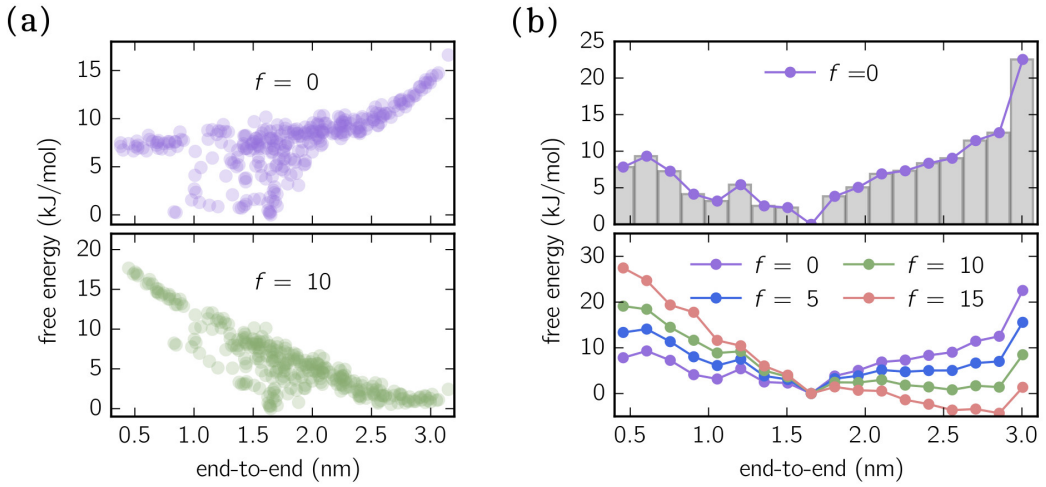

Figure 6.3 – **Free energy landscape of deca-alanine.** (a) Scatter plot of free energy against end-to-end distance of every individual centroid for $f = 0$ kJ/mol·nm (upper panel) and $f = 10$ kJ/mol·nm (lower panel). (b) The free energy landscape integrated along the end-to-end distance (bin size = 0.15 nm) is shown for the unbiased dynamics (upper panel) and for all ensembles (lower panel) $f = \{0, 5, 10, 15\}$ kJ/mol·nm.

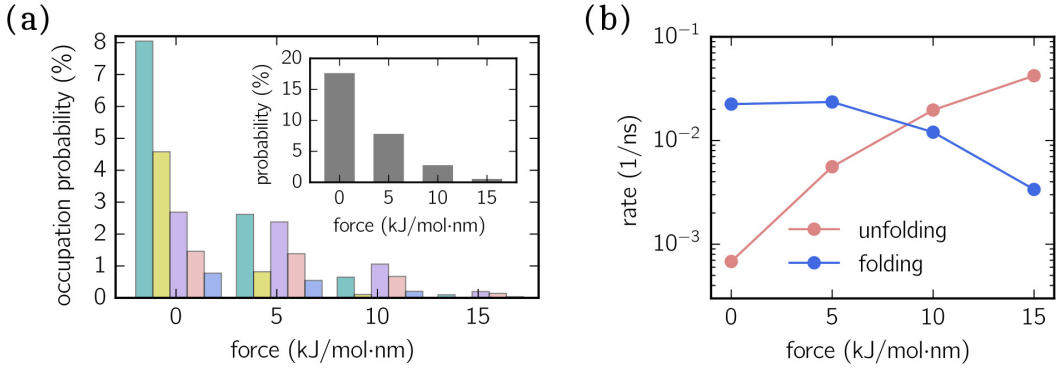

Figure 6.4 – **Folding/unfolding rates and misfolded states of deca-alanine.** (a) Occupation probability of misfolded configurations for biased/unbiased ensembles. Colors match the metastable sets shown in Figure 6.2. The inset shows the total percentage to occupy a misfolded state. (b) Inverse mean-first passage times between stretched (light blue) and correctly folded (green) states for all constant-force ensembles.

Finally, folding/unfolding rates of deca-alanine are determined by computing inverse mean-first passage times (MFPTs) between correctly folded (green) and stretched (light blue) config-

urations. Figure 6.4(b) shows both rates for all ensembles ranging from $7 \times 10^{-4} - 4 \times 10^{-2}$ ns^{-1} (unfolding) and from $2 \times 10^{-2} - 3 \times 10^{-3}$ ns^{-1} (folding).

6.3 Example II: Beta-alanine

For the second example we choose the beta-peptide β-HALA$_8$, which is an octamer containing eight alanine residues. The biasing force is applied along the vector connecting the C_β atoms of the first and last residue, which we refer to as the end-to-end distance of the peptide. The peptide is described by the GROMOS 53A6 force field [99]. Periodic boundary conditions are imposed on the cubic simulation box (7 nm box length) which is filled with ≃12000 water molecules. We use the SPC water model [14] and the GROMOS 53A6 force field in order to make our results comparable to previous studies [137]. We create 48 trajectories from different starting configurations accumulating 2–2.5 μs of the constant-force ensembles $f = \{0, 2, 5, 10\}$ kJ/mol·nm.

As an initial representation of the configuration space we select all C_β distances excluding the two closest neighbors, as well as all hydrogen bond forming donor-acceptor distances that stabilize the beta-helical structure. More precisely, the distance between C=O of the ith residue and N–H of the $(i-2)$th residue. This 21-dimensional space is reduced by TICA to three effective dimensions (cumulative kinetic variance > 95%). Including the end-to-end distance, a reduced four-dimensional configuration space representation is then discretized into 300 micro-states employing the k-means clustering algorithm. To obtain the MSMs for different field strengths, we apply the TRAM estimator for a lag time of 0.6 ns.

6.3.1 Configuration space

To visualize the reduced configuration space of beta-alanine, we show in Figure 6.5(a) subspaces spanned by the first two TICA components and the first TICA component with the end-to-end distance. Circles represent centroid positions with different colors distinguishing different metastable sets. Metastable sets are again identified by applying the PCCA+ algorithm to the unbiased transition matrix. Overall, seven different metastable sets are found distinguishing correctly folded beta-helical structures (green), stretched configurations (light blue), misfolded structures (cyan and red) and folding intermediates. Structures characterized as folding intermediates are exclusively coiled along the centered residue (purple), coiled at the N-terminus (blue) and unfolded but weakly stabilized by one or two hydrogen bounds (brown), which also stabilize the correctly folded beta-helix. Similar to the configuration of deca-alanine, stretched configurations are separated from unfolded configurations according to their end-to-end distance.

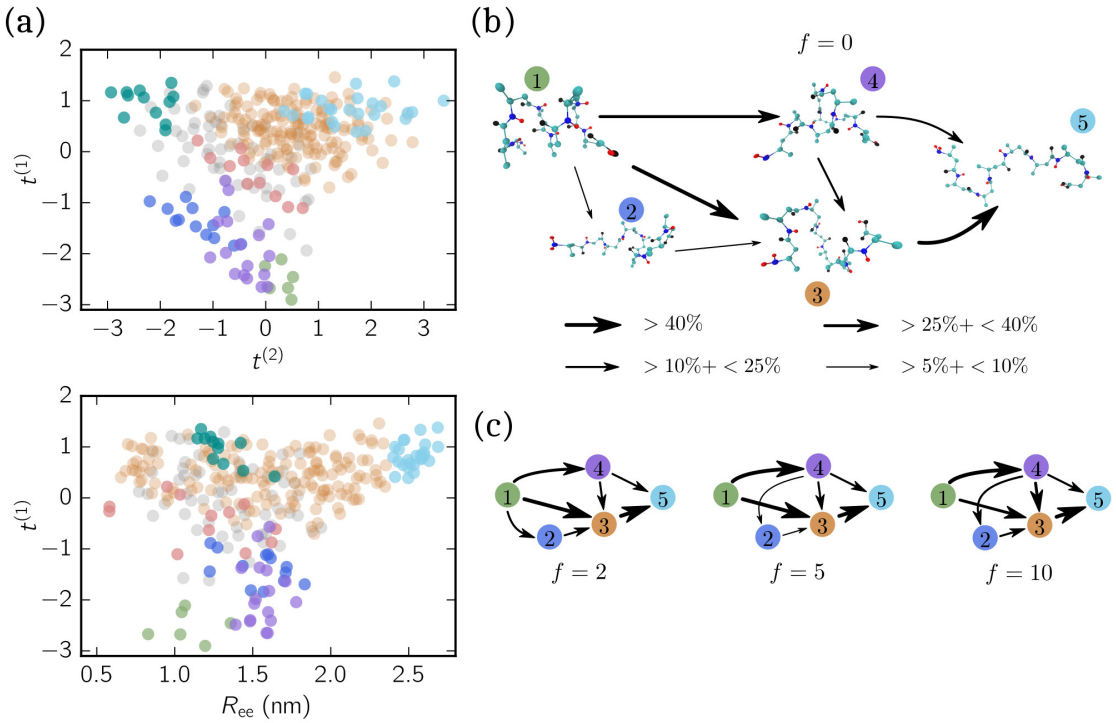

Figure 6.5 – **Configuration space of beta-alanine.** (a) Configuration space projection of the first and second TICA component (upper panel), as well as the first TICA component and end-to-end distance (lower panel) for $f = 0$ kJ/mol·nm. Colored circles indicate centroids belonging to different metastable sets. Gray circles illustrate transition states that do not belong to any metastable set. (b) Reactive transition network of unfolding pathways for the unbiased dynamics. Each snapshot represents its metastable set [cf. (a)] while arrows indicate the direction of net transitions and their thickness the transition frequency. (c) The same reactive transition network as in (b) but for $f = \{5, 10, 15\}$ kJ/mol · nm.

6.3.2 Free energy landscape and unfolding pathways

In Figure 6.6(a), we show the free energy of every individual centroid plotted against the end-to-end distance for the unbiased ensemble. Compared with the previous example [see Figure 6.3(a)], the free energy landscape shows a rather broad minimal region between $R_{ee} = 1 - 1.5$ nm coinciding with six out of seven metastable sets [cf. Figure 6.5(a)]. However, when selecting the first TICA component, the free energy landscape exhibits three distinguishable minima as shown in Figure 6.6(b) (upper panel) for which we have integrated out the remaining configuration space components. The first minimum for $t^{(1)} \leq -2.5$ belongs to correctly folded structures. The second minimum at $t^{(1)} \simeq -1.5$ we identify as folding intermediates (purple and blue colored structures) and finally the global minimum at $t^{(1)} \simeq 0.5$ corresponds to (brown) unfolded structures. Note that colors match the colored sets illustrated in Figure 6.5(a). The free energy projection along the first TICA component, thus, reveals that the beta-helical structure is thermodynamically unstable, which is hidden when only taking the end-to-end distance into account.

The lower panel of Figure 6.5(b) shows the same analysis but for all ensembles. Interestingly, for $f = \{0,2,5\}$ kJ/mol·nm all free energy landscapes are almost identical but for $f =$ 10 kJ/mol·nm a noticeable change occurs. The free energy landscape between the first two minima becomes rather flat, while a new minimum appears at $t^{(1)} \simeq -0.7$ and the global minimum is shifted to $t^{(1)} \simeq 0.7$.

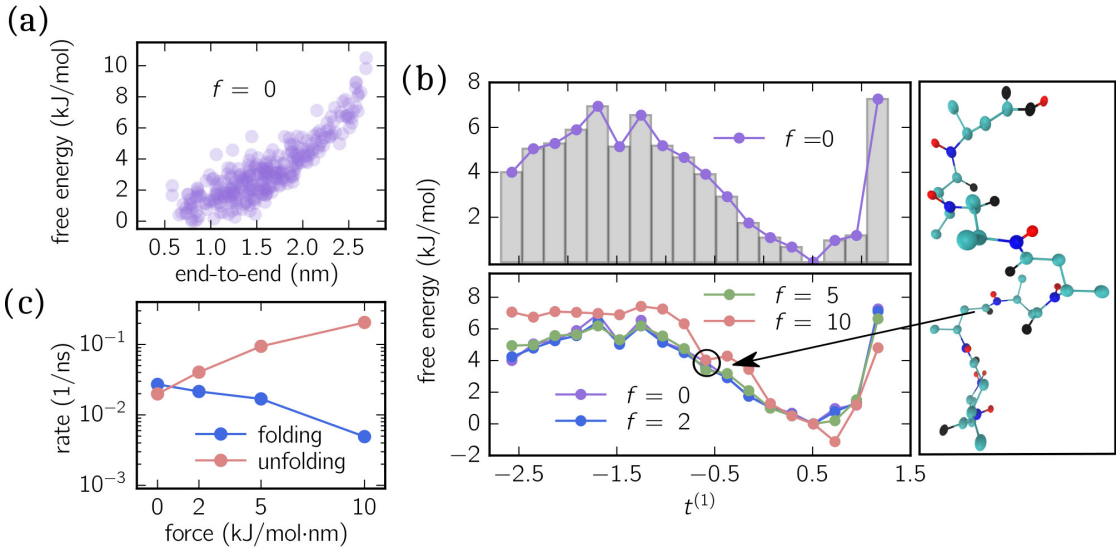

Figure 6.6 – **Free energies and unfolding/folding rates of beta-alanine.** (a) Scatter plot of free energies against the end-to-end distance for $f = 0$ kJ/mol·nm. (b) Free energy landscape projected and integrated along the first TICA component for $f = 0$ kJ/mol·nm (upper panel) and for all ensembles $f = \{0,2,5,10\}$ kJ/mol·nm (lower panel). All free energy landscapes are adjusted with respect to the minimum for $f = 0$ kJ/mol·nm. An exemplary snapshot is shown representing structures that correspond to the circled point of the free energy landscape for $f = 10$ kJ/mol·nm. (c) Inverse MFPTs between correctly folded (green) and stretched (light blue) configurations for all ensembles.

To obtain quantitative unfolding/folding rates, inverse MFPTs are computed connecting folded and stretched configurations. Both rates are displayed for all constant-force ensembles in Figure 6.5(c), ranging from $2.7 \times 10^{-2} - 5 \times 10^{-3}$/ns (unfolding) and from $2 \times 10^{-2} - 2 \times 10^{-1}$/ns (folding).

Next, we quantify the importance of folding intermediates by determining unfolding pathways of the peptide. To this end, we employ **discrete transition path theory** [106] identifying and weighting pathways starting from folded and ending with stretched configurations. In Figure 6.5(b), we show the transition network for the unbiased ensemble, where metastable sets are represented by an exemplary snapshot and enumerated by colored labels with colors matching the colors used in Figure 6.5(a). The arrows point in the direction of the reactive flux with their width indicating the relative importance of its transition (the thicker the more important). For example, the transition from state 1 to 4 illustrates that between 10% and 25% of all transition pathways – connecting folded and stretched configurations – run along

this transition. The main contribution is given by pathways containing folding intermediates, while structures coiled at the N-terminus (dark blue) are less important.

Interestingly, pathways including misfolded configurations do not contribute (< 5%) to the unfolding pathways indicating that these states act as kinetic traps. A similar picture is obtained when applying the same analysis to the constant-force ensembles shown in Figure 6.5(c), although slight differences appear. For $f = \{5, 10\}$ kJ/mol·nm, for instance, the transition from $1 \rightarrow 2$ vanishes. However, a new transition from $4 \rightarrow 1$ appears or rather becomes more dominant. Moreover, the most dominant pathway changes from $1 \rightarrow 3 \rightarrow 5$ for $f = \{0, 2\}$ kJ/mol·nm to $1 \rightarrow 4 \rightarrow 3 \rightarrow 5$ for $f = \{5, 10\}$ kJ/mol·nm. Numerical values for all shown transitions are presented in Table 6.1.

Table 6.1 – Weights of transition pathways for the unfolding process of beta-alanine for all constant-force ensembles. Transition labels match labels shown in Figure 6.5(b) and (c).

# transition	$f = 0$	$f = 2$	$f = 5$	$f = 10$
$1 \rightarrow 2$	8%	10%	-	-
$1 \rightarrow 3$	45%	42%	46%	32%
$1 \rightarrow 4$	31%	34%	42%	54%
$2 \rightarrow 3$	8%	10%	7%	11%
$3 \rightarrow 5$	74%	72%	74%	68%
$4 \rightarrow 2$	-	-	8%	10%
$4 \rightarrow 3$	20%	22%	23%	26%
$4 \rightarrow 5$	10%	11%	10%	18%

6.4 Discussion

6.4.1 Choice of order parameter

For both examples we showed that the end-to-end distance is not a well-suited order parameter since many metastable sets cannot be distinguished solely by the end-to-end distance [cf. right panel of Figure 6.2 and lower panel of Figure 6.5(a)] requiring an orthogonal order parameter. In the study of Hazel *et al.* [57], the authors come to a similar conclusion. In particular, they investigate the folding dynamics for deca-alanine in vacuum and water, where a second order parameter describing the alpha-helical content is employed yielding better insight into the folding/unfolding pathways. Here we employed TICA as a semi-automated routine creating a minimal set of orthogonal order parameters being able to accurately describe the configuration space. Besides alpha-helical/beta-helical structure, the chosen TICA space is able to resolves misfolded configurations, see upper panels of Figure 6.2 and Figure 6.5(a).

6.4.2 Free energy landscapes and metastable configurations

To estimate the transition times between two different types of states, e.g. stretched and folded configurations, a commonly used method is to determine the **potential of mean force** (PMF) by umbrella sampling methods [115, 101]. Therefore, many umbrella potentials are placed along one or multiple order parameters. Using more than one order parameter becomes quickly very expensive, making it very difficult to accurately determine a PMF when a "good" order parameter is not known beforehand or does simply not exist. However, the more complex (larger) a peptide/protein is, the higher the order parameter space becomes if essential energy barriers should not be missed. Furthermore, we want to emphasize that umbrella sampling along only the end-to-end distance is prone to errors as it is easy to miss certain metastable sets. For example, in Figure 6.2 the hairpin (yellow) and wrongly coiled (cyan) structures share a common end-to-end distance. The barrier between both states is, however, very large because the hairpin structure has to be extended first.

To compare the unbiased PMF of deca-alanine with previous studies, we compute the PMF along the end-to-end distance [see Figure 6.3(b)]. While its global minimum is found at 1.46 nm in ref. [57] and at around 1.4 nm in ref. [69], we find it at 1.6 nm which agrees well when regarding the different definitions of the end-to-end distance. The general form of the PMF is in very good agreement with ref. [69], but differs from ref. [57] due to the different force field used. More precisely, the authors employed the CHARMM36 force field which destabilizes the alpha-helical structure.

The PMF along the first TICA component of beta-alanine [cf. Figure 6.6(b)] reveals three stable regions corresponding to folded structures, an intermediate and weakly folded structures with the latter being the most dominant region. When only considering the end-to-end distance [see Figure 6.6(a)], the integrated PMF decreases monotonically (data not shown) in agreement to the results obtained by Uribe *et al.* [137]. Moreover, in their study the authors employed force-probe MD simulations separating the N- and C-terminus with constant velocity. By means of force-extension curves and hydrogen bond analyses, the authors conclude that the unfolding pathway includes two intermediates before obtaining a fully stretched peptide. We identify these intermediates with the purple and brown colored sets shown in Figure 6.5(a). These findings are further supported by the reactive flux analysis as illustrated in Figure 6.5(b) and (c). Imposing a constant force shifts the weight of transitions pathways. Specifically, the transition $1 \to 4$ becomes more frequent while the transition $1 \to 3$ becomes less frequent with increasing bias. In other words, when pulling stronger it is more likely to pass purple colored states than without pulling.

The additional minimum in the PMF for $f = 10\,\mathrm{kJ/mol{\cdot}nm}$ [lower panel Figure 6.6(b)], however, suggests that the free energy landscape is drastically modified by the biasing force. To elucidate the mechanism causing the additional minimum, we visualize an exemplary configuration for $t^{(1)} \simeq -0.7$ exhibiting a low free energy. The snapshot shown in Figure 6.6(b) illustrates that the outer residues are stretched while the inner residues are stabilized by hydrogen bonds.

Outlook

In this thesis we have extended the concepts of Markov state modeling to systems that are driven into a nonequilibrium steady state. Albeit most soft matter systems are, up to today, simulated in thermal equilibrium, many real-life applications require nonequilibrium conditions. Of particular importance are biomolecular systems which physiological functions depend typically on their native environment, especially on the chemical potential of chemical reactants. Therefore, an accurate description of systems involving chemical reactions requires a molecular chemiostat. The chemoistat keeps the chemical potential constant for molecules that are consumed or produced during a chemical reaction, and thus drives the system into a nonequilibrium steady state.

From the computational point of view, to include chemical reactions in MD simulations requires either a quantum mechanical treatment or the use of reactive force fields [138]. While a quantum mechanical treatment becomes very quickly numerically inaccessible for large systems, reactive force fields have not been well developed yet, in particular for biomolecular applications. However, a first step in the direction of incorporating chemical reactions in large scale soft matter simulations, even if somehow artificial, can be found in the work of Zeravic *et al.* [151]. There, the authors show that the self-replication of colloidal clusters is significantly accelerated when supplying the system with colloidal dimers acting as catalysts. Moreover, the authors identify an effective reaction (replication) network containing multiple production cycles, which are analogous to the probability current cycles identified in nonequilibrium Markov state models.

Other applications of nonequilibrium Markov state modeling involve systems including hydrodynamics flow fields, e.g., shear flow. One could, for instance, think of flow fields that can accelerate the self-assembly of microscopic structures. In this context, it has been shown that, depending on the imposed strain rate, shear flow can enhance or suppress the crystallization of hard spheres [110].

Furthermore, the construction of NE-MSMs for periodically driven systems can be of value not only for understanding how a molecular dynamics is influenced by externally applied oscillating fields but for steering molecular systems via systematically designing periodic protocols. For example one could use an oscillating electromagnetic field to increase or decrease the population of specific molecular conformations.

Finally, combining Markov state modeling with externally biased systems (time-dependent or time-independent) appears to be a promising strategy for linking experiments and simulations, in particular of biomolecular systems, even with the computational resources currently available.

In any case, all these different types of applications have in common that they cover time scales spanning many orders of magnitude. Therefore, we are convinced that nonequilibrium Markov state modeling will be a valuable tool not just to bridge the large spectrum of time scales present in many driven soft matter systems but also to offer meaningful insight into their nonequilibrium dynamics.

Bibliography

[1] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.

[2] C. C. Aggarwal and C. K. Reddy. *Data clustering: algorithms and applications*. CRC Press, 2014.

[3] B. J. Alder and T. E. Wainwright. Studies in molecular dynamics. i. general method. *J. Chem. Phys.*, 31(2):459–466, 1959.

[4] A. Alexander-Katz, M. F. Schneider, S. W. Schneider, A. Wixforth, and R. R. Netz. Shear-flow-induced unfolding of polymeric globules. *Phys. Rev. Lett.*, 97(13):138101, 2006.

[5] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids*. Oxford university press, 1989.

[6] B. Altaner. *Foundations of Stochastic Thermodynamics*. PhD thesis, Universität Göttingen, 2015.

[7] B. Altaner, S. Grosskinsky, S. Herminghaus, L. Katthän, M. Timme, and J. Vollmer. Network representations of nonequilibrium steady states: Cycle decompositions, symmetries, and dominant paths. *Phys. Rev. E*, 85:041133, 2012.

[8] B. Altaner and J. Vollmer. Fluctuation-preserving coarse graining for biochemical systems. *Phys. Rev. Lett.*, 108(22):228101, 2012.

[9] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72(4):2384–2393, 1980.

[10] S. Arrhenius. On the influence of carbonic acid in the air upon the temperature of the ground. *Philos. Mag.*, 41(251):237–276, 1896.

[11] C. L. Asbury and G. Van Den Engh. Trapping of DNA in nonuniform oscillating electric fields. *Biophys. J.*, 74(2):1024–1030, 1998.

[12] R. D. Astumian. Stochastic conformational pumping: a mechanism for free-energy transduction by molecules. *Annu. Rev. Biophys.*, 40:289–313, 2011.

[13] A. C. Barato and U. Seifert. Thermodynamic uncertainty relation for biomolecular processes. *Phys. Rev. Lett.*, 114:158101, 2015.

[14] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91(24):6269–6271, 1987.

[15] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81(8):3684–3690, 1984.

[16] L. Berthier and G. Biroli. Theoretical perspective on the glass transition and amorphous materials. *Rev. Mod. Phys.*, 83:587–645, 2011.

[17] R. B. Best and G. Hummer. Protein folding kinetics under force from molecular simulation. *J. Am. Chem. Soc*, 130(12):3706–3707, 2008.

[18] K. Binder. Applications of Monte Carlo methods to statistical physics. *Rep. Prog. Phys.*, 60(5):487, 1997.

[19] V. Böhmer. Calixarenes, macrocycles with (almost) unlimited possibilities. *Angew. Chem. Int. Ed.*, 34(7):713–745, 1995.

[20] G. R. Bowman and V. S. Pande. Protein folded states are kinetic hubs. *Proc. Natl. Acad. Sci. USA*, 107(24):10890–10895, 2010.

[21] G. R. Bowman, V. S. Pande, and F. Noé. *An introduction to Markov state models and their application to long timescale molecular simulation*, volume 797. Springer Science & Business Media, 2013.

[22] O. Braun, A. Hanke, and U. Seifert. Probing molecular free energy landscapes by periodic loading. *Phys. Rev. Lett.*, 93:158105, Oct 2004.

[23] N.-V. Buchete and G. Hummer. Coarse master equations for peptide folding dynamics. *J. Phys. Chem. B*, 112(19):6057–6069, 2008.

[24] A. Budi, F. S. Legge, H. Treutlein, and I. Yarovsky. Electric field effects on insulin chain-b conformation. *J. Phys. Chem. B*, 109(47):22641–22648, 2005.

[25] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys*, 126(1):014101, 2007.

[26] H. B. Callen. *Thermodynamics and an Introduction to Thermostatistics*. Wiley-VCH, 2. edition, 1985.

[27] M. Carrion-Vazquez, A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, and J. M. Fernandez. Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. USA*, 96(7):3694–3699, 1999.

[28] C. Cecconi, E. A. Shank, C. Bustamante, and S. Marqusee. Direct observation of the three-state folding of a single protein molecule. *Science*, 309(5743):2057–2060, 2005.

[29] J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *J. Chem. Phys*, 126(15):155101, 2007.

[30] J. D. Chodera, W. C. Swope, J. W. Pitera, and K. A. Dill. Long-time protein folding dynamics from short-time molecular dynamics simulations. *Multiscale Model Simul.*, 5(4):1214–1226, 2006.

[31] M. Cieplak, M. Henkel, J. Karbowski, and J. R. Banavar. Master equation approach to protein folding and kinetic traps. *Phys. Rev. Lett.*, 80:3654–3657, Apr 1998.

[32] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante. Verification of the crooks fluctuation theorem and recovery of rna folding free energies. *Nature*, 437:231, 2005.

[33] N. D. Conrad, R. Banisch, and C. Schütte. Modularity of directed networks: Cycle decomposition approach. *J. Comput. Dyn.*, 2:1–24, 2015.

[34] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60:2721–2726, Sep 1999.

[35] L. F. Cugliandolo. Course 7: Dynamics of glassy systems. In *Slow Relaxations and nonequilibrium dynamics in condensed matter*, pages 367–521. Springer, 2003.

[36] T. Darden, D. York, and L. Pedersen. Particle mesh ewald: An $N \cdot \log N$ method for ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092, 1993.

[37] P. G. De Gennes. Coil-stretch transition of dilute flexible polymers under ultrahigh velocity gradients. *J. Chem. Phys*, 60(12):5030–5042, 1974.

[38] C. A. F. de Oliveira, D. Hamelberg, and J. A. McCammon. Estimating kinetic rates from accelerated molecular dynamics simulations: Alanine dipeptide in explicit solvent as a case study. *J. Chem. Phys*, 127(17), 2007.

[39] R. Delgado-Buscalioni. Cyclic motion of a grafted polymer under shear flow. *Phys. Rev. Lett.*, 96(8):088303, 2006.

[40] P. S. Doyle, B. Ladoux, and J.-L. Viovy. Dynamics of a tethered polymer in shear flow. *Phys. Rev. Lett.*, 84:4769–4772, 2000.

[41] J. L. England. Statistical physics of self-replication. *J. Chem. Phys*, 139(12):121923, 2013.

[42] N. J. English, G. Y. Solomentsev, and P. O'Brien. Nonequilibrium molecular dynamics study of electric and low-frequency microwave fields on hen egg white lysozyme. *J. Chem. Phys.*, 131(3):07B614, 2009.

[43] M. Esposito, U. Harbola, and S. Mukamel. Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems. *Rev. Mod. Phys.*, 81:1665–1702, Dec 2009.

[44] M. Esposito and G. Schaller. Stochastic thermodynamics for "maxwell demon" feedbacks. *Europhys. Lett.*, 99(3):30003, 2012.

[45] M. Esposito and C. Van den Broeck. Second law and landauer principle far from equilibrium. *Europhys. Lett.*, 95(4):40004, 2011.

[46] J. Esque and M. Cecchini. Accurate calculation of conformational free energy differences in explicit water: The confinement–solvation free energy approach. *J. Phys. Chem. B*, 119(16):5194–5207, 2015.

[47] M. G. Evans and M. Polanyi. Some applications of the transition state method to the calculation of reaction velocities, especially in solution. *Trans. Faraday Soc.*, 31:875–894, 1935.

[48] H. Eyring. The activated complex and the absolute rate of chemical reactions. *Chemical Reviews*, 17(1):65–77, 1935.

[49] S. Fortunato. Community detection in graphs. *Physics reports*, 486(3):75–174, 2010.

[50] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*, volume 1. Academic press, 2001.

[51] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Rev. Mod. Phys*, 70(1):223, 1998.

[52] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England. Dissipation bounds all steady-state current fluctuations. *Phys. Rev. Lett.*, 116:120601, 2016.

[53] H. Grabert. *Projection operator techniques in nonequilibrium statistical mechanics*. Springer Berlin Heidelberg, 1982.

[54] H. Grubmüller. Force probe molecular dynamics simulations. *Protein-Ligand Interactions: Methods and Applications*, pages 493–515, 2005.

[55] H. Grubmüller, B. Heymann, and P. Tavan. Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*, (5251):997–999, 1996.

[56] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after kramers. *Rev. Mod. Phys.*, 62(2):251, 1990.

[57] A. Hazel, C. Chipot, and J. C. Gumbart. Thermodynamics of deca-alanine folding in water. *J. Chem. Theory Comput.*, 10(7):2836–2844, 2014.

[58] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M Fraaije. Lincs: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, 18(12):1463–1472, 1997.

[59] C. Hijón, P. Español, E. Vanden-Eijnden, and R. Delgado-Buscalioni. Mori–zwanzig formalism as a practical computational tool. *Farad. Discuss.*, 144:301–322, 2010.

[60] T. L. Hill. *Free energy transduction and biochemical cycle kinetics*. Courier Corporation, 2004.

[61] W. G. Hoover. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):1695, 1985.

[62] J. Howard. *Mechanics of Motor Proteins and the Cytoskeleton*. Sinauer Associates, Publishers, 2001.

[63] G. Hummer and A. Szabo. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Acad. Sci. U.S.A.*, 98:3658, 2001.

[64] R. B. Israel, J. S. Rosenthal, and J. Z. Wei. Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings. *Math. Finance*, 11(2):245–265, 2001.

[65] B. Isralewitz, M. Gao, and K. Schulten. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.*, 11(2):224–230, 2001.

[66] M. Janke, Y. Rudzevich, O. Molokanova, T. Metzroth, I. Mey, G. Diezemann, P. E. Marszalek, J. Gauss, V. Böhmer, and A. Janshoff. Mechanically interlocked calix[4]arene dimers display reversible bond breakage under force. *Nat. Nanotechnol.*, 4(4):225–229, 2009.

[67] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, Apr 1997.

[68] D.-Q. Jiang, M. Qian, and M.-P. Qian. *Mathematical theory of nonequilibrium steady states: on the frontier of probability and dynamical systems*. Springer, 2003.

[69] S. Jo, D. Suh, Z. He, C. Chipot, and B. Roux. Leveraging the information from Markov state models to improve the convergence of umbrella sampling simulations. *J. Phys. Chem. B*, 120(33):8733–8742, 2016.

[70] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.

[71] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*, 79(2):926–935, 1983.

[72] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, 1996.

[73] S. L. Kalpazidou. *Cycle representations of Markov processes*, volume 28. Springer Science & Business Media, 2007.

[74] S. Katz, J. L. Lebowitz, and H. Spohn. Phase transitions in stationary nonequilibrium states of model lattice systems. *Phys. Rev. B*, 28:1655–1658, Aug 1983.

[75] M. S. Z Kellermayer, S. B. Smith, H. L. Granzier, and C. Bustamante. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science*, 276(5315):1112–1116, 1997.

[76] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[77] J. Kubelka, J. Hofrichter, and W. A. Eaton. The protein folding "speed limit". *Curr. Opin. Struct. Biol.*, 14(1):76–88, 2004.

[78] P. A. Kuchment. *Floquet theory for partial differential equations*. Springer, 2013.

[79] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comput. Chem.*, 13(8):1011–1021, 1992.

[80] D. P. Landau and K. Binder. *A guide to Monte Carlo simulations in statistical physics*. Cambridge university press, 2014.

[81] R. G. Larson. The rheology of dilute solutions of flexible polymers: Progress and problems. *J. Rheol.*, 49(1):1–70, 2005.

[82] A. S. Lemak, N. K. Balabaev, Y. N. Karnet, and Y. G. Yanovsky. The effect of a solid wall on polymer chain behavior under shear flow. *J. Chem. Phys*, 108(2):797–806, 1998.

[83] S. Liepelt and R. Lipowsky. Kinesin's network of chemomechanical motor cycles. *Phys. Rev. Lett.*, 98:258102, Jun 2007.

[84] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw. How fast-folding proteins fold. *Science*, 334(6055):517–520, 2011.

[85] H. Lu, B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.*, 75(2):662–671, 1998.

[86] A. D. MacKerell, N. Banavali, and N. Foloppe. Development and current status of the charmm force field for nucleic acids. *Biopolymers*, 56(4):257–265, 2000.

[87] R. T. McGibbon and V. S. Pande. Efficient maximum likelihood parameterization of continuous-time Markov processes. *J. Chem. Phys*, 143(3):034109, 2015.

[88] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

[89] P. Metzner, E. Dittmer, T. Jahnke, and C. Schütte. Generator estimation of Markov jump processes. *J. Comput. Phys.*, 227(1):353–375, 2007.

[90] A. S. J. S. Mey, H. Wu, and F. Noé. xtram: Estimating equilibrium expectations from time-correlated simulation data at multiple thermodynamic states. *Phys. Rev. X*, 4(4):041018, 2014.

[91] R. K. Mishra, G. Mishra, D. Giri, and S. Kumar. Scaling of hysteresis loop of interacting polymers under a periodic force. *J. Chem. Phys.*, 138(24):244905, 2013.

[92] F. Müller-Plathe. Coarse-graining in polymer simulation: from the atomistic to the mesoscopic scale and back. *ChemPhysChem*, 3(9):754–769, 2002.

[93] K. C. Neuman and A. Nagy. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature methods*, 5(6):491, 2008.

[94] M. Newman. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA*, 103(23):8577–8582, 2006.

[95] M. Newman. *Networks: An Introduction*. Oxford university press, 2010.

[96] A. Nicholls, K. A. Sharp, and B. Honig. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct., Funct., Bioinf.*, 11(4):281–296, 1991.

[97] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. Weikl. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. USA*, 106(45):19011, 2009.

[98] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, 81(1):511–519, 1984.

[99] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6. *J. Comput. Chem.*, 25(13):1656–1676, 2004.

[100] E. Paci and M. Karplus. Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. *J. Mol. Biol.*, 288(3):441–459, 1999.

[101] S. Park and K. Schulten. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.*, 120(13):5946–5961, 2004.

[102] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys*, 52(12):7182–7190, 1981.

[103] G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys*, 139(1):015102, 2013.

[104] N. Plattner and F. Noé. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun*, 6, 2015.

[105] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.*, 117(1):1–19, 1995.

[106] J.-H. Prinz, B. Keller, and F. Noé. Probing molecular kinetics with Markov models: metastable states, transition pathways and spectroscopic observables. *Phys. Chem. Chem. Phys.*, 13(38):16912–16927, 2011.

[107] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys*, 134(17):174105, 2011.

[108] A. Puglisi, S. Pigolotti, L. Rondoni, and A. Vulpiani. Entropy production and coarse graining in Markov processes. *J. Stat. Mech.*, page P05015, 2010.

[109] A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev.*, 136:A405–A411, 1964.

[110] D. Richard and T. Speck. The role of shear in crystallization kinetics: From suppression to enhancement. *Sci. Rep*, 5, 2015.

[111] F. Rico, L. Gonzalez, I. Casuso, M. Puig-Vidal, and S. Scheuring. High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science*, 342(6159):741–743, 2013.

[112] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. Reversible unfolding of individual titin immunoglobulin domains by afm. *Science*, 276(5315):1109–1112, 1997.

[113] M. Rief, F. Oesterhelt, B. Heymann, and H. E. Gaub. Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science*, 275(5304):1295–1297, 1997.

[114] S. Röblitz and M. Weber. Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification. *Adv. Data. Anal. Classif.*, 7(2):147–179, 2013.

[115] B. Roux. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Com.*, 91(1-3):275–282, 1995.

[116] M. Sarich and C. Schütte. Utilizing hitting times for finding metastable sets in non-reversible Markov chains. Technical Report 14-32, ZIB, Takustr. 7, 14195 Berlin, 2014.

[117] M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz, and F. Noé. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.*, 11:5525–5542, October 2015.

[118] T. Schlesier and G. Diezemann. Performance of different force fields in force probe simulations. *J. Phys. Chem. B*, 117(6):1862–1871, 2013.

[119] T. Schlesier, T. Metzroth, A. Janshoff, J. Gauss, and G. Diezemann. Reversible hydrogen bond network dynamics: Molecular dynamics simulations of calix[4]arene-catenanes. *J. Phys. Chem. B*, 115(20):6445–6454, 2011.

[120] M. Schliwa and G. Woehlke. Molecular motors. *Nature*, 422(6933):759–765, 2003.

[121] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.*, 48(4):571, 1976.

[122] S. Schuler, T. Speck, C. Tietz, J. Wrachtrup, and U. Seifert. Experimental test of the fluctuation theorem for a driven two-level system with time-dependent rates. *Phys. Rev. Lett.*, 94(18):180602, 2005.

[123] C. Schütte, W. Huisinga, and P. Deuflhard. Transfer operator approach to conformational dynamics in biomolecular systems. In *Ergodic theory, analysis, and efficient simulation of dynamical systems*, pages 191–223. Springer, 2001.

[124] U. Seifert. Stochastic thermodynamics: principles and perspectives. *Eur. Phys. J. B*, 64(3):423–431, 2008.

[125] U. Seifert. Stochastic thermodynamics, fluctuation theorems, and molecular machines. *Rep. Prog. Phys.*, 75:126001, 2012.

[126] Udo Seifert. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.*, 95:040602, Jul 2005.

[127] D. E. Smith, H. P. Babcock, and S. Chu. Single-polymer dynamics in steady shear flow. *Science*, 283(5408):1724–1727, 1999.

[128] M. Sotomayor and K. Schulten. Single-molecule experiments in vitro and in silico. *Science*, 316(5828):1144–1148, 2007.

[129] T. Speck, J. Mehl, and U. Seifert. Role of external flow and frame invariance in stochastic thermodynamics. *Phys. Rev. Lett.*, 100(17):178302, 2008.

[130] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.*, 76(1):637–649, 1982.

[131] W. C. Swope, J. W. Pitera, and F. Suits. Describing protein folding kinetics by molecular dynamics simulations. 1. Theory. *J. Phys. Chem. B*, 108(21):6571–6581, 2004.

[132] W. C. Swope, J. W. Pitera, F. Suits, M. Pitman, M. Eleftheriou, B. G. Fitch, R. S. Germain, A. Rayshubski, T. J. C. Ward, Y. Zhestkov, and R. Zhou. Describing protein folding kinetics by molecular dynamics simulations. 2. Example applications to alanine dipeptide and a β-hairpin peptide. *J. Phys. Chem. B*, 108(21):6582–6594, 2004.

[133] V. Tozzini. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.*, 15(2):144–150, 2005.

[134] B. Trendelkamp-Schroer and F. Noé. Efficient bayesian estimation of Markov model transition matrices with given stationary distribution. *J. Chem. Phys.*, 138(16):164113, 2013.

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

[136] B. Trendelkamp-Schroer, H. Wu, F. Paul, and F. Noé. Estimation and uncertainty of reversible Markov models. *J. Chem. Phys*, 143(17):174101, 2015.

[137] L. Uribe, J. Gauss, and G. Diezemann. Comparative study of the mechanical unfolding pathways of α-and β-peptides. *J. Phy. Chem. B*, 119(26):8313–8320, 2015.

[138] A. C. T Van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard. Reaxff: a reactive force field for hydrocarbons. *J. Phys. Chem. A*, 105(41):9396–9409, 2001.

[139] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry*, volume 1. Elsevier, 3. edition, 2007.

[140] C. Velez-Vega, E. E. Borrero, and F. A. Escobedo. Kinetics and reaction coordinate for the isomerization of alanine dipeptide by a forward flux sampling protocol. *J. Chem. Phys*, 130(22), 2009.

[141] V. A. Voelz, G. R. Bowman, K. Beauchamp, and V. S. Pande. Molecular simulation of ab initio protein folding for a millisecond folder ntl9 (1- 39). *J. Am. Chem. Soc*, 132(5):1526–1528, 2010.

[142] L. Volkmann. *Fundamente der Graphentheorie*. Springer Lehrbuch Mathematik. Springer-Verlag, 1996.

[143] G. A. Voth. *Coarse-graining of condensed phase and biomolecular systems*. CRC press, 2008.

[144] H. Wang and C. Schütte. Building Markov state models for periodically driven non-equilibrium systems. *J. Chem. Theory Comput.*, 11(4):1819–1831, 2015.

[145] H. Wang, C. Schütte, G. Ciccotti, and L. Delle Site. Exploring the conformational dynamics of alanine dipeptide in solution subjected to an external electric field: A nonequilibrium molecular dynamics simulation. *J. Chem. Theory Comput.*, 10(4):1376–1386, 2014.

[146] W. Wang and Y. Zhang. On fuzzy cluster validity indices. *Fuzzy Sets Syst.*, 158(19):2095–2117, 2007.

[147] J. K. Weber, D. Shukla, and V. S. Pande. Heat dissipation guides activation in signaling proteins. *Proc. Natl. Acad. Sci. USA*, 112(33):10377–10382, 2015.

[148] R. G. Winkler. Semiflexible polymers in shear flow. *Phys. Rev. Lett.*, 97:128301, 2006.

[149] H. Wu, F. Paul, C. Wehmeyer, and F. Noé. Multiensemble Markov models of molecular thermodynamics and kinetics. *Proc. Natl. Acad. Sci. USA*, 113(23):E3221–E3230, 2016.

[150] Z. T. Yew, M. Schlierf, M. Rief, and E. Paci. Direct evidence of the multidimensionality of the free-energy landscapes of proteins revealed by mechanical probes. *Phys. Rev. E*, 81:031923, Mar 2010.

[151] Z. Zeravcic and M. P. Brenner. Self-replicating colloidal clusters. *Proc. Natl. Acad. Sci. USA*, 111(5):1748–1753, 2014.

[152] Y. Zhang, A. Donev, T. Weisgraber, B. J. Alder, M. D. Graham, and J. J. de Pablo. Tethered DNA dynamics in shear flow. *J. Chem. Phys.*, 130(23):234902, 2009.

[153] R. Zwanzig. *Nonequilibrium statistical mechanics*. Oxford University Press, 2001.

Colophon. This thesis was typeset with LaTeX using the EPFL template available at http://phd.epfl.ch/thesistemplates. It uses the Utopia font created by C. Dellagiacoma, which is a combination of OpenSource Utopia by Adobe and Greek characters provided by P. Pıchaureau within the OpenSource MathDesign package for LaTeX.

List of Publications

Some parts of this thesis are adaptations and/or extensions of the following publications and manuscripts:

- **F. Knoch** and T. Speck. Cycle-based nonequilibrium Markov state modeling for periodic driving. *manuscript in preparation*.
- **F. Knoch**, K. Schäfer, G. Diezemann and T. Speck. Dynamic coarse graining allows to fill the gap between atomistic simulations and experimental investigations of mechanical unfolding. *preprint arXiv:1710.09699*.
- **F. Knoch** and T. Speck. Unfolding dynamics of small peptides biased by constant mechanical forces. *Mol. Syst. Des. Eng.*, 2017.
- **F. Knoch** and T. Speck. Nonequilibrium Markov state modeling of the globule-stretch transition. *Phys. Rev. E*, 95(1):012503, 2017.
- **F. Knoch** and T. Speck. Cycle representatives for the coarse-graining of systems driven into a non-equilibrium steady state. *New J. Phys.*, 17(11):115004, 2015.

Other publications of the author published during or prior to the PhD program:

- V. Wulf, **F. Knoch**, T. Speck and C. Sönnichsen. Gold nanorods as plasmonic sensors for particle diffusion. *J. Phys. Chem. Lett.*, 7(23):4951–4955, 2016.
- J. Rother, C. Richter, L. Turco, **F. Knoch**, I. Mey, S. Luther, A. Janshoff, E. Bodenschatz and M. Tarantola. Crosstalk of cardiomyocytes and fibroblasts in co-cultures. *Open Biol.*, 5(6):150038, 2015.
- E. Böhm, C. Pflumm, H. Heil, S. Meyer, **F. Knoch**, L.-I. Rodriguez, B. Burkhart, F. Eckes, K. Stegmaier and H. Buchholz. Blue fluorescent OLEDs for printed display applications. *Proceedings IDW 14*, OLED4-1, 2014.
- **F. Knoch**, M. Tarantola, E. Bodenschatz and W.-J. Rappel. Modeling self-organized spatio-temporal patterns of PIP3 and PTEN during spontaneous cell polarization. *Phys. Biol.*, 11(4):046002, 2014.
- **F. Knoch**, D. Morozov, M. Boggio-Pasqua and G. Groenhof. Steering the excited state dynamics of a photoactive yellow protein chromophore analogue with external electric fields. *Comp. Theor. Chem.*, 1040:120–125, 2014.