Chapter 1

Background

1.1 Statistical Mechanics

Statistical Mechanics is a branch of physics that deals with complex systems with many degrees of freedom using statistics and probability theory. It emerged from thermodynamics that deals with heat and temperature and is based on three laws that were defined by observing nature. While this is a macroscopic theory treating temperature and heat as natural quantities, statistical mechanics dives into the microscopic world, resulting in a deeper understanding of the thermodynamic principles. Statistical Physics derives macroscopic matter properties through inference from its microscopic features and marginalising details that are not of importance. [8].

1.1.1 Density of States

To understand this view on thermodynamics, we have to define a *microscopic* and *macroscopic* state. We assume a set of N particles in a box, each particle labelled i has a precise position x_i and momentum p_i. The microscopic state of a system contains all these information. Mathematically it is a $6N$-dimensional vector in the *phase space* that contains all possible microstates of a system. The phase space grows large even for simple systems of a few particles. Fortunately, the overall information in the system is of minor importance, compared to macroscopic variables, like the

system's energy E, its volume V and others [8].

Each macroscopic system consists of a large number of possible microscopic states, and it is impossible to determine in which of these states the system is. We define the number of possible states depending on a macroscopic variable as the *density of states* $\Omega(E, V, ..)$.

1.1.2 Ensemble Averages

We prepare a system and wait for some time such that the system relaxes to *equilibrium*. From a thermodynamical point of view, this means that no heat Q is transferred to or from the system. From a microscopic perspective, we have to assume that the heat can fluctuate but no heat is transferred on average and the distribution of microstates does not change. Only the microstate of the system does change in time. Hypothetically, one can record all these microstates to find their distribution. Similarly, one may duplicate the equilibrated system with the same macroscopic variables and record the microscopic configuration for each system individually. Both methods give the same answer if the system is *ergodic*. That means that the system can evolve in time and reach any point in phase space with given macroscopic constraints by only passing through allowed states. This can be problematic if, for instance, a constant energy prevents the system from crossing an energy barrier and discovering new states satisfying the constraint. The outcome of the measurement would depend on its initial state. For this reason, we will assume an *ensemble* of the system, i.e. the duplication with equal macroscopic constraints [8].

In practice, it is inconvenient to count microstates in order to find the distribution of microstates. There are three frequently used ensembles in equilibrium statistical mechanics with known probability distribution [20]:

- The microcanonical ensemble of N particles is isolated from the environment and does not exchange energy or particles with it. Each microstate occurs with the probability $p = \frac{1}{N}$.

- The canonical ensemble, where the system is connected weakly to a heat bath of constant temperature and interchanges energy with it. The probability distribution is $p_i = \frac{1}{Z} \exp\left(-\frac{E_i}{k_B T}\right)$, where Z is the normalisation, k_B is the Boltzmann constant and T is the temperature of the reservoir and the system.

- The grandcanonical ensemble is connected weakly to a reservoir and interchanges particles and energy with it. The probability distribution is $p_i = \frac{1}{Y} \exp\left(-\frac{E_i - \mu N}{k_B T}\right)$, where Y is the normalisation and μ is the chemical potential.

The aforementioned probability distributions will not be derived here because they emerge directly from Jaynes Maximum Caliber as presented in the corresponding chapter **??**.

Given the probability of all microstates in the system, we could calculate the ensemble average of an observable O by

$$\langle O \rangle = \sum_{i}^{\text{microstates}} p_i O_i. \tag{1.1}$$

For calculating the probability distribution with respect to the macroscopic observable O, we have to sum over all probabilities of all microstates satisfying the observable

$$p(O) = \sum_{i \in O} p_i. \tag{1.2}$$

We note that it is useful to define the density of states according to the controlling parameter of the ensemble, e.g. the energy in the canonical ensemble. In this case we can write

$$p(O) = \sum_{E} \frac{\Omega(E, O)}{Z} \exp\left(-\frac{E}{k_B T}\right). \tag{1.3}$$

These basic relations open up a way to reweight microstates between ensembles. We assume an energy distribution $p_T(E)$ in canonical ensemble at temperature T sampled from a simulation of a system. Given above relation, we can reweight the distribution to another temperature T'' by

$$
\begin{aligned}
p_{T''}(E) &= \frac{\Omega(E)}{Z} \exp\left(-\frac{E}{k_B T''}\right) \\
&= \frac{\Omega(E)}{Z'} \exp\left(-\frac{E}{k_B T} + \frac{E}{k_B T'}\right) \\
&= \frac{p_T(E)}{Z'} \exp\left(-\frac{E}{k_B T'}\right),
\end{aligned}
\tag{1.4}
$$

using the reference data $p_T(E)$. T' is the shift from the original temperature T to reach the target temperature T''. This technique will be derived and discussed using Jaynes Maximum Caliber in chapter **??**. It relies on knowing the density of states. In fact, the density of states is a starting point to calculate all thermodynamic properties of the system. Therefore, it is the quantity that is of most interest when running a simulation [11].

1.1.3   Molecular Dynamics

Simulations are frequently used to discover the phase space of complex systems. Molecular Dynamics (MD) simulations collect information on the micro- and macrostates of the system by numerically solving Newton's equation of motion forward in time. It is thus of importance to set up the simulation such that the ensemble of choice is sampled correctly. For instance, sampling the canonical ensemble requires to model the heat exchange with the reservoir correctly. Heat is transferred via collision between particles of the system and the reservoir. We do not know when the collision occur so they are modelled using a stochastic variable $\zeta(t)$ acting on the particles in collision. The reservoir and the system are at the same temperature, meaning that each degree of freedom carries the same amount of energy on average by the equipartition theorem [20].

$$m\langle v(t)^2\rangle = k_{\mathrm{B}}T. \tag{1.5}$$

The average energy transferred to the system should be zero, so the stochastic variable has zero mean. Furthermore, weak coupling ensures that collision are uncorrelated and we write

$$\langle \zeta(t)\rangle = 0$$
$$\langle \zeta_i(t)\zeta_j(t')\rangle = C\delta(t - t')\delta_{ij}, \tag{1.6}$$

where C is a constant to be determined and the indices i, j denote single components of a vector. The equation of motion to be solved becomes

$$m\ddot{\mathbf{x}} = -m\gamma\dot{\mathbf{x}} + \boldsymbol{\zeta}(t), \tag{1.7}$$

where m is the mass of a particle and γ is a constant friction term that dissipates energy. Integrating this equation once gives

$$\dot{\mathbf{x}}(t) = \mathbf{v_0}e^{-\gamma t} + \int_0^t e^{\gamma s}\boldsymbol{\zeta}(\boldsymbol{s})\,\mathrm{d}s\,, \tag{1.8}$$

4

where $\mathbf{v_0}$ is the initial velocity. The average over an ensemble is

$$\langle \dot{\mathbf{x}}(t) \rangle = \mathbf{v_0} e^{-\gamma t}. \tag{1.9}$$

The fluctuation around the average are

$$\langle (\mathbf{v}(t) - \mathbf{v_0} e^{-\gamma t})^2 \rangle = \frac{C}{m^2 \gamma}(1 - e^{-2\gamma t}). \tag{1.10}$$

Letting the system relax to equilibrium for large t gives the equilibrium fluctuation. Equating these fluctuations with the equipartition theorem in equation 1.5 gives

$$C = m\gamma k_{\mathrm{B}}T, \tag{1.11}$$

allowing to find the strength of the random force. This ensures fluctuations characteristic for the chosen temperature and sampling the correct ensemble [17]. Note that this derivation uses the fluctuation around the equilibrium value so the thermalisation is not bound to work in off-equilibrium.

The presented Langevin thermostat is only one possible way of controlling the system, other methods are the DPD thermostat [25] or the Nosé-Hoover thermostat [10] that can be applied in various equilibrium ensembles. The Langevin equation of the system of interest is numerically integrated using the velocity-verlet algorithm [30] or the Runge-Kutta method [12] for overdamped dynamics.

1.2   Stochastic Thermodynamics

Stochastic thermodynamics is a framework for soft-matter systems under general off-equilibrium conditions: (i) The system is embedded in a liquid solution functioning as a heat-bath with constant temperature T, (ii) the system is pushed out of equilibrium with external mechanical forces or unbalanced chemical potentials and (iii) fluctuations play a major role in the system [22]. These assumptions were experimentally verified beyond the linear regime on particles moved by laser traps [1] and DNA or RNA manipulated by optical tweezers [14]. The main plan is to adapt ideas like heat, entropy, energy and their relations from classical thermodynamics and translate them to the nano- and microscopic world without assuming equilibrium. We start with the Langevin equation for overdamped motion

$$\dot{\mathbf{x}} = \mu \mathbf{F}(\mathbf{x}, \lambda) + \boldsymbol{\zeta}(t), \tag{1.12}$$

Technical Point 1.1 Stratonovich vs Itô Integration

The overdamped Langevin Equation 1.12 can be written in a more general form

$$\mathrm{d}x = A(x,t) + B(x,t)\,\mathrm{d}W(t)\,, \tag{1.13}$$

where $A(x,t) = \frac{\mu}{m}F(x)$, $B(x,t) = \frac{1}{m}$ and $\mathrm{d}W(t) = \zeta(t)\,\mathrm{d}t$ being the Wiener process [29]. The process is reduced to one dimension for simplicity. The Wiener process is not a continuous function by definition because it integrates over a stochastic variable $W(t)$. Applying Riemann integration to the stochastic term shows

$$\lim_{n\to\infty} \sum_{j}^{n} B(x,\tau_j)(W(t_{j+1}) - W(t)), \tag{1.14}$$

where τ_j lies in $[t_j, t_{j+1}]$. The limit does not exist because W(t) depends on a stochastic variable. The difference in Itô and Stratonovich integration lies in the choice

$$\begin{aligned} \tau_j &= t_j && \text{(Itô)} \\ \tau_j &= \frac{t_j + t_{j+1}}{2} && \text{(Stratonovich)}, \end{aligned} \tag{1.15}$$

resulting in different solutions of the integral. The different choices in τ_j indicate if the δ-like collision occurs in the beginning or during the discrete time-step. For example, Itô integration is chosen for modelling photon emission or chemical reaction, Stratonovich integration for external effects like connected noise generators [29]. The latter one is chosen in the present case because the modelled degrees of freedom are much slower than the time between collision. The large amount of collisions imitates the behaviour of noise generators.

where the force $\mathbf{F}(\mathbf{x}, \lambda) = \frac{\partial V(x,\lambda)}{\partial \mathbf{x}} + \mathbf{f}(x, \lambda)$ consists of a conservative and non-conservative part, $\boldsymbol{\zeta}(t)$ is Gaussian noise and μ is the mobility. The thermal noise is assumed to be independent of the external disturbance and the temperature of the thermal bath is given by the equilibrium Einstein relation $T = \frac{D}{\mu k_{\mathrm{B}}}$, where k_{B} is the Boltzmann constant and D is the diffusion coefficient. The parameter λ controls an external change of the potential or forces on a system and introduces time-dependent off-equilibrium effects [22].

Initialising the system at some point $\mathbf{x_0}$ and integrating forward in time produces a possible trajectory. We have introduced ensembles for microstates and ensemble averages in the previous section 1.1.2. We will now do this for the off-equilibrium ensem-

ble defined above. Dynamics of off-equilibrium processes are often time-dependent, so microstates are not sufficient for description and we have to collect trajectories. Through Stratonovich integration (see Technical Point 1.1) of equation 1.12, we find for the weight of a trajectory $\mathbf{x}(t)$ with starting point $\mathbf{x_0}$ [23]

$$p[\mathbf{x}(t), \lambda(t)|\mathbf{x_0}] = \prod_i \exp\left[-\int_{-t}^{t} \mathrm{d}\tau \, \frac{(\dot{x}_i - \mu F_i(\mathbf{x}, \lambda))^2}{4D} + \frac{\mu}{2}\frac{\partial F_i(\mathbf{x}, \lambda)}{\partial x_i}\right], \qquad (1.16)$$

where i iterates over all dimension. The integration over all trajectories denoted by $\mathrm{d}[\mathbf{x}(t)]$ with initial distribution $p(\mathbf{x_0})$ should be normalised

$$1 = \int \int \mathrm{d}\mathbf{x_0} \, \mathrm{d}[\mathbf{x}(t)] \, p[\mathbf{x}(t), \lambda(t)|\mathbf{x_0}]p(\mathbf{x_0}). \qquad (1.17)$$

Similarly, one can define the ensemble expectation value over trajectories

$$\langle O \rangle = \int \int \mathrm{d}\mathbf{x_0} \, \mathrm{d}[\mathbf{x}(t)] \, O[\mathbf{x}(t), \lambda(t)]p[\mathbf{x}(t), \lambda(t)|\mathbf{x_0}]p(\mathbf{x_0}). \qquad (1.18)$$

The action is defined equivalent to equilibrium statistical mechanics by $p[\mathbf{x}(t), \lambda(t)] = \exp(-\mathcal{A}[\mathbf{x}(t), \lambda(t)])$ so one finds from equation 1.16

$$\mathcal{A}[\mathbf{x}(t), \lambda(t)] = \sum_i \int_{-t}^{t} \mathrm{d}\tau \, \frac{(\dot{x}_i - \mu F_i(\mathbf{x}, \lambda))^2}{4D} + \frac{\mu}{2}\frac{\partial F_i(\mathbf{x}, \lambda)}{\partial x_i}. \qquad (1.19)$$

The existence of a trajectory $\mathbf{x}(t)$ implies the existence of the same trajectory under time reversal $\tilde{\mathbf{x}}(-t)$ [3]. Time reversal includes a reversed control protocol $\lambda(t) \to \tilde{\lambda}(-t)$ and trajectories $\mathbf{x}(t) \to \tilde{\mathbf{x}}(-t)$, $\dot{\mathbf{x}}(t) \to \dot{\tilde{\mathbf{x}}}(-t)$. The action of the reversed trajectory becomes

$$\mathcal{A}[\tilde{\mathbf{x}}(t), \lambda(t)] = \sum_i \int_{-t}^{t} \mathrm{d}\tau \, \frac{\dot{x}_i^2 + 2\mu F_i(\mathbf{x}, \lambda)\dot{x}_i + F_i(\mathbf{x}, \lambda)^2}{4D} + \frac{\mu}{2}\frac{\partial F_i(\mathbf{x}, \lambda)}{\partial x_i}$$
$$= \mathcal{A}[\mathbf{x}(t), \lambda(t)] + \int_{-t}^{t} \mathrm{d}\tau \, \frac{1}{T}\mathbf{F}(\mathbf{x}, \lambda) \cdot \dot{\mathbf{x}}. \qquad (1.20)$$

The microscopic definition of heat dissipated $q[\mathbf{x}(t), \lambda(t)]$ along a trajectory can be calculated by the ratio of the weight of a forward trajectory and its time inverted counterpart. The entropy produced in the reservoir $s_m[\mathbf{x}(t), \lambda(t)]$ is related to the

heat by dividing with the temperature like in thermodynamics.

$$\frac{p[\mathbf{x}(t), \lambda(t)|\mathbf{x}_0]}{p[\tilde{\mathbf{x}}(t), \tilde{\lambda}(t)|\tilde{\mathbf{x}}_0]} = \exp\left(\frac{1}{T}\int_{-t}^{t} d\tau\, \mathbf{F}(\mathbf{x}, \lambda) \cdot \dot{\mathbf{x}}\right)$$

$$= \exp\left(\frac{q[\mathbf{x}(t), \lambda(t)]}{T}\right) \quad (1.21)$$

$$= \exp\left(s_m[\mathbf{x}(t), \lambda(t)]\right).$$

The more heat is generated in a forward process, the less likely is the reverse process [22]. This relates forward and time-inverted processes for non-equilibrium processes. Note that this exponential relation implies $s_m[\mathbf{x}(t), \lambda(t)] = -s_m[\tilde{\mathbf{x}}(t), \tilde{\lambda}(t)]$ and therefore, that the heat produced in a forward process is the same as the heat consumed by its time-inverted process and vice versa.

We expect stochastic thermodynamics to obey the main laws of thermodynamics. The first law states that energy is conserved in an isolated system. A change in internal energy ΔU consists of heat Q transferred to the system and the work W done by the system on its environment

$$\Delta U = Q - W. \quad (1.22)$$

for a Langevin process on a microscopic level, this becomes [24]

$$d\omega = du + dq$$

$$= \left(\frac{\partial V}{\partial \lambda} \dot{\lambda}\, d\tau\right) + f\, dx\,, \quad (1.23)$$

where du is identified with $\frac{\partial V}{\partial \lambda}\dot{\lambda}\, d\tau$ as the change in potential at fixed particle position and dq is identified by $f\, dx$ as the heat dissipated by the medium due to the non-conservative force f in agreement to the finding in equation 1.21. The second law of thermodynamics states that two initially isolated and equilibrated systems increase their combined entropy when they get in contact. The entropy does not change when the systems were already in equilibrium with each other. The second law is validated by defining the total entropy production

$$\Delta s_{\text{tot}}[\mathbf{x}(t), \lambda(t), p_{\text{i}}, p_{\text{f}}] = \log\frac{p[\mathbf{x}(t)|\mathbf{x}_0]p_{\text{i}}(\mathbf{x}_0)}{p[\tilde{\mathbf{x}}(t)|\tilde{\mathbf{x}}_0]p_{\text{f}}(\tilde{\mathbf{x}}_0)}$$

$$= \log\frac{p[\mathbf{x}(t)|\mathbf{x}_0]}{p[\tilde{\mathbf{x}}(t)|\tilde{\mathbf{x}}_0]} + \log\frac{p_{\text{i}}(\mathbf{x}_0)}{p_{\text{f}}(\tilde{\mathbf{x}}_0)} \quad (1.24)$$

$$= s_m[\mathbf{x}(t), \lambda(t)] + s_{\text{sys}}(p_i, p_f),$$

8

where the distribution change along the trajectory from initial p_i to final p_f is the system contribution to total entropy production. $s_m[\mathbf{x}(t), \lambda(t)]$ is the contribution of the reservoir. The result is used to derive

$$\langle \exp\left(-\Delta s_\text{tot}\right)\rangle = \int \int \mathrm{d}\mathbf{x_0}\, \mathrm{d}[\mathbf{x}(t)]\, p_\text{i}(\mathbf{x_0})\exp\left(-\Delta s_\text{tot}\right)$$
$$= \int \int \mathrm{d}\tilde{\mathbf{x}}_\mathbf{0}\, \mathrm{d}[\tilde{\mathbf{x}}(t)]\, p[\tilde{\mathbf{x}}(t)|\tilde{\mathbf{x}}_\mathbf{0}]p_\text{f}(\tilde{\mathbf{x}}_\mathbf{0}) \qquad (1.25)$$
$$= 1.$$

The second law of thermodynamics $\langle \Delta s_\text{tot}\rangle \geq 0$ follows directly from $\langle \exp -\Delta s_\text{tot}\rangle = 1$. Note that equation 1.25 does not exclude single trajectories to violate the second law, but these trajectories are exponentially suppressed and the second law is only valid for ensembles of trajectories. The contributions to entropy production are well separated in contributions of the system and medium, like the first law in equation 1.23 [22].

1.2.1   Transfer Operator

The Langevin equation 1.7 can produce numerous realisations of a process due to its dependence on a stochastic variable. All of these realisations are possible pathways and are combined in an ensemble of trajectories. A probability distribution $P(x)$ can be constructed depending on the space passing point x [2]. This distribution is evolved in time according to a generator $\mathscr{L}$,containing all dynamical information by knowledge of a all trajectories passing x, denoted by

$$P(x_{t+\delta t}) = \mathscr{L} \circ P(x_t). \qquad (1.26)$$

The initial probability distribution can be chosen according to physical conditions. The infinitesimal generator $\mathscr{L}$ and dependence on the whole history $P(x_t)$ is inconvenient for computational purposes and is approximated by a time-discretised transfer operator.

1.2.2   Markov Processes

We want the transfer operator to evolve the system over a time-range $\tau \gg \delta t$ and we want it to fulfil the *Markov property*. The probability distribution $P(y_i, t_i)$ of a

Markov process only depends on the information at the previous time-point t_{i-1}. For n successive time-points $(t_1 < t_2 < .. < t_n)$ we call

$$P(t_n, y_n | t_1, y_1; t_2, y_2; ...; t_{n-1}, y_{n-1}) = P(t_n, y_n | t_{n-1}, y_{n-1}) \tag{1.27}$$

the time dependent Markovian *transition probability*. Note that the time evolution of a probability distribution can be expressed using the Markov property by the hierarchy

$$P(t_1, y_1; t_2, y_2; t_3, y_3) = P(t_1, y_1) P(t_2, y_2 | t_1, y_1) P(t_3, y_3 | t_2, y_2), \tag{1.28}$$

where $P(t_1, y_1)$ is the initial probability distribution. The Markov process is fully characterised by knowledge of the transition probabilities and the initial probability distribution. Integrating over y_2 and dividing by $P(t_1, y_1)$ gives

$$P(t_1, y_1 | t_3, y_3) = \int \mathrm{d}y_2 P(t_2, y_2 | t_1, y_1) P(t_3, y_3 | t_2, y_2) \tag{1.29}$$

where the definition of conditional probabilities was used on the left-hand side [29]. Equation 1.29 is known as the Chapman-Kolmogorov equation and will be used to test if the Markov property is fulfilled.

The transfer operator with the Markov property is defined by

$$P_{t+\tau}(y) = \mathscr{T}(\tau) \circ P_t(x) = \int_\Omega \mathrm{d}y W(x|y); \tau) P_t(y), \tag{1.30}$$

where Ω represents the full conformational space of the system and $W(x|y; \tau)$ is a transition probability in continuous space from state y to state x after time-step τ. The operator is related to the time-continuous operator in equation 1.26 by

$$\mathscr{T}(\tau) = \exp\left(\mathscr{L}\tau\right), \tag{1.31}$$

where the eigenvectors Ψ_i of both are the same and the eigenvalue λ_i of $\mathscr{T}$ and Λ_i of $\mathscr{L}$ are related by $\lambda_i = \exp(\tau\Lambda_i)$ [2].

We note that the stationary distribution is reached when $\mathscr{T}(\tau) \circ P_s(x) = P_s(x)$, i.e. it is the eigenvector with eigenvalue 1 of the transfer operator [19]. It is assumed to be singular such that the stationary distribution of a system is unique. The smaller eigenvalues are related to the relaxation dynamics of the system, as will be discussed in detail in section 1.3.

10

1.2.3   Non-Equilibrium Steady States

A Non-Equilibrium Steady State (NESS) is a special type of non-equilibrium process because the fluxes and population densities do not depend on time. It is generated by driving a system with time-independent non-conservative forces for a long time. The resulting stationary distribution and fluxes are generally unknown. The external driving results in an input heat flow in the system. The same amount of heat is dissipated in a coupled reservoir to maintain the steady state [5]. The contribution to total entropy production in equation 1.24 of the system is zero and only the medium produces entropy.

Assuming a Markovian process with transition probability $W(y|y')$ from state y' to y per unit time, we may describe the dynamics by the Master equation [29]

$$\frac{\partial P(y,t)}{\partial t} = \int \mathrm{d}y' \left(W(y|y')P(y',t) - W(y'|y)P(y,t) \right). \tag{1.32}$$

The first term on the right hand side of the equation adds up all probability fluxes to state y, while the second term adds up all fluxes going out of that state. The result of all probability fluxes is the change in population of that state. We want to deal with NESS, so the time dependence of the distribution is set to $\frac{\partial P(y,t)}{\partial t} = 0$. Under this condition, the equation can be solved by assuming the fluxes cancel pairwise

$$W(y|y')P(y',t) = W(y'|y)P(y,t). \tag{1.33}$$

In fact, this is the *detailed balance* condition that holds true in equilibrium systems. However, a general solution can be found by use of the normalisation $\int \mathrm{d}y'\, W(y'|y) = 1$

$$P(y) = \int \mathrm{d}y'\, W(y|y')P(y',t), \tag{1.34}$$

called the *global balance* condition that always holds true for NESS. It shows that the probability to be in a state y is the sum over all probability fluxes from all states. The fluxes transport probability in cycles in the system, unlike for detailed balance where fluxes cancel each other. Note that detailed balance implies global balance but not the other way around [29].

We can introduce another NESS condition by integrating the microscopic reversibility condition in equation 1.21. The dependence on an external controlling parameter is dropped in a NESS because the disturbance is permanent. For future use, we want

11

to discretise the space. The indicator function assigns each point in space p to a state

$$\chi_i(p) = \begin{cases} 1 & \text{if} \quad p \in i \\ 0 & \text{if} \quad p \notin i. \end{cases} \qquad (1.35)$$

Each discrete state will be called *microstate* [2], not to be confused with the microstate of statistical physics described in 1.1 that consists of all coordinates of every particle in the system. The present microstates emerges from coarse-graining the conformational space. The coarse-grained states can be chosen according to the system studies. In the next step, the probability distribution over entropy production ΔS for a given transition of microstates is sought. We integrate over all existing trajectories with entropy production ΔS that start in microstate i and end in microstate j. This produces a trajectory ensemble average over all time forward processes, denoted by subscript F:

$$P_{ij}(\Delta S) = \int \delta(\chi_i(\mathbf{x}_0)) \, \mathrm{d}\mathbf{x}_0 \int \delta(\chi_j(\mathbf{x}_1)) \, \mathrm{d}\mathbf{x}_1$$
$$\int \mathrm{d}[\mathbf{x}(t)] \, P(\mathbf{x}_0) p[\mathbf{x}(+t)] \delta(\Delta S - \Delta S_F[\mathbf{x}(t)]) \qquad (1.36)$$

where x_0 and x_1 denotes the start and endpoint of a trajectory. Making use of equation 1.21 transforms the expression to the inverse trajectory ensemble, denoted by subscript R:

$$P_{ij}(\Delta S) = \int \delta(\chi_i(\mathbf{x}_0)) \, \mathrm{d}\mathbf{x}_0 \int \delta(\chi_j(\mathbf{x}_1)) \, \mathrm{d}\mathbf{x}_1$$
$$\int \mathrm{d}[\mathbf{x}(t)] \, P(x_0) p[\tilde{\mathbf{x}}(-t)|\tilde{\lambda}(-t)] e^{\Delta S} \delta(\Delta S + \Delta S_R) \qquad (1.37)$$
$$= e^{\Delta S} P_{ji}(-\Delta S)$$

where the relation $\Delta S_F = -\Delta S_R$ is used, meaning that the heat gained from the reservoir in a forward trajectory is the same heat ejected for the time-reversed trajectory. Note, that the probability distributions $P(\mathbf{x}_0) = P(\mathbf{x}_1)$ remain the system for a non-equilibrium steady state. The resulting equation $\frac{P_{ij}}{P_{ji}} = \exp(\Delta S_{ij})$ relates the amount of heat dissipated or drawn from the reservoir with the probability of the forward and backward jump between predefined microstates, acting like a coarse-grained version of equation 1.21. In fact, it is a special case of Crooks fluctuation theorem [4] with defined starting and end state in long time approximation. It simplifies to detailed balance in equilibrium when Boltzmann statistics can be assumed.

1.2.4   Driven Particle in a periodic Potential

The following construction of a *Markov State Model* will be presented on the example of a driven single particle. The model is a simple case of a NESS [26] and consists of a particle driven by a non-conservative force f that may emerge from magnetic fields, mechanical flows, or mechanical dragging. The particle is influenced by a potential shown in figure 1.1. An analogous experiment was performed using silica spheres on a tilted crystalline glass substrate [15]. Another discrete example is a weakly interacting Bose-Einstein condensate on a one-dimensional optical lattice [13]. The overdamped equation of motion for the particle is given by

$$0 = -\frac{\partial U(x)}{\partial x} - \gamma\dot{x} + \sqrt{2\gamma k_\mathrm{B}T}\zeta(t) + f, \tag{1.38}$$

where a reservoir of temperature T is coupled to the system by friction constant γ and $\zeta(t)$ is a δ-correlated Gaussian process with mean 0. Results are presented in reduced units, where the box size is set to $\mathcal{L}$, the mass of the particle is set to $\mathcal{M}$, and energy is measured in ϵ. The temperature is $T = 1\,\epsilon/k_\mathrm{B}$, the energy barriers shown in figure 1.1 are varied between $2 - 4\,\epsilon$ and the unit of time is $\mathcal{T} = \mathcal{L}\sqrt{\mathcal{M}/\epsilon}$. The integration time step was set to $\delta t = 10^{-5}\,\mathcal{T}$. The non-conservative force is varied between 0 and $9\,\epsilon/\mathcal{L}$ and the space is separated in 60 microstates of same size.

The presented potential is described by:

$$\begin{aligned}
U(x) = & U_0\left(\frac{1}{1 + \exp(-2k(x - s_{N+1}))}\right) \\
& - U_N\left(\frac{1}{1 + \exp(-2k(x - s_0))}\right) \\
& + \sum_{j=0}^{j\leq N} U_j\left(\frac{1}{1 + \exp(-2k(x - s_j))} - \frac{1}{1 + \exp(-2k(x - s_{j+1}))}\right),
\end{aligned} \tag{1.39}$$

where s_j shifts each summand of the potential to its chosen position, U_j denotes the potential of the N local maxima or minima. The potential has periodicity p such that $U(x) = U(x + p)$ and $s_{N+1} = s_0 + p$. Parameter k describes the smoothness of the potential and is chosen as $k = -N\log\left(\frac{1}{0.999} - 1\right)$. k should be chosen larger to minimise a force discontinuity at the periodic boundary. The potential becomes a set of Heaviside steps functions without discontinuity for $k \to \infty$. This potential allows us to construct a model with arbitrary barrier heights, well depths, steepness of transition regions, and number and position of wells and barriers.
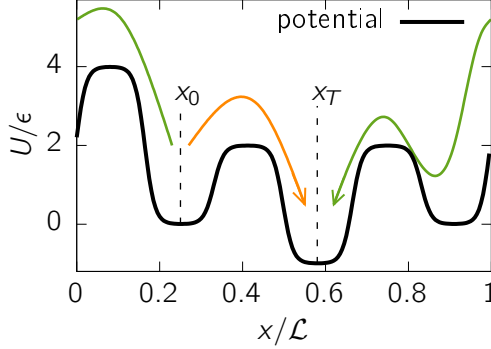
13

Figure 1.1: An example potential, where the minima or chosen to be at position $s_{\min} = \frac{3}{12}\mathcal{L}, \frac{7}{12}\mathcal{L}, \frac{11}{12}\mathcal{L}$ with potential minima $U_{\min} = 0\,\epsilon, -1\,\epsilon, 0\,\epsilon$ respectively and maxima are at position $s_{\max} = \frac{1}{12}\mathcal{L}, \frac{5}{12}\mathcal{L}, \frac{9}{12}\mathcal{L}$ with potential $U_{\max} = 4\,\epsilon, 2\,\epsilon, 2\,\epsilon$ respectively. The green and orange arrow represent two different set of paths going from position x_i to x_j.

Entropy Production

We want to derive an expression for the local entropy productions ΔS_{ij} as introduced in 1.2.3. The local entropy production of a single continuous trajectory $\mathbf{x}(t)$ is given by

$$\Delta S[x(t)] = \int \mathrm{d}t\, \frac{\mathbf{F} \cdot \dot{\mathbf{x}}}{k_{\mathrm{B}}T}, \tag{1.40}$$

where $\dot{\mathbf{x}}$ is the velocity and T is the temperature [21]. Making use of numerically discretised trajectories from simulation, $\mathbf{x}(t) \approx \{\mathbf{x}_k\}$, ΔS is approximated between starting and target points x_0 and x_T, respectively,

$$\Delta S[\{x_k\}] \approx \frac{1}{2T} \sum_{d} \sum_{t=1}^{t=T} \left(x_t^{(d)} - x_{t-1}^{(d)} \right) \left(F^{(d)}(\mathbf{x}_t) + F^{(d)}(\mathbf{x}_{t-1}) \right), \tag{1.41}$$

where Stratonovich integration is used (see technical point 1.1) and d iterates over the dimensions.

The solution above requires to integrate along trajectories and average over a trajectory ensemble. We approximate the solution by ignoring the fluctuations, allowing us to apply Riemann integration. We make use of the analytic expression

14

$\mathbf{F} = -\frac{\partial U(=x)}{\partial \mathbf{x}} + \mathbf{f}$ and find

$$\begin{aligned}
\Delta S[x(t)] &= \frac{1}{k_\mathrm{B}T} \int \mathrm{d}t \sum_d \left(\frac{\partial U}{\partial x_d} \frac{\partial x_d}{\partial t} + f_d \frac{\partial x_d}{\partial t} \right) \\
&= \frac{1}{k_\mathrm{B}T} \int \mathrm{d}t \frac{\mathrm{d}U}{\mathrm{d}t} + \sum_d \left(\int \mathrm{d}t \, f_d \frac{\partial x_d}{\partial t} \right) \\
&= \frac{U(x_T) - U(x_0)}{k_\mathrm{B}T} + \sum_d \left(\int \mathrm{d}t \, f_d \frac{\partial x_d}{\partial t} \right).
\end{aligned} \tag{1.42}$$

The non-conservative force $\mathbf{f}$ cannot be expressed by a potential and the integral depends on the path of $\mathbf{x}$. For the current 1D-case, two different sets of paths exist between x_0 and x_T by using the periodic boundaries as indicated in figure 1.1. For equilibrium systems with $f = 0$ both pathways have the same entropy production, otherwise there are two different solutions. By choosing the time-length of trajectories short, the longer path has negligible weight and we can assume a unique solution to the entropy production. The expression for the local entropy production becomes

$$\Delta S(x_0, x_T) \approx \frac{U(x_T) - U(x_0) + (x_T - x_0)f}{k_\mathrm{B}T}, \tag{1.43}$$

only depending on start and endpoint of the trajectory. The solution of this equation differs from the solution of equation 1.41 by ignoring fluctuations and giving us an analytic estimate for the entropy production. The path ensemble average of the entropy production is assumed to agree with the analytic solution as discussed in section **??**. The approximated analytic solution is used throughout the thesis as it is much easier to solve.

1.3 Markov State Model

Markov State Modeling (MSM) aims to map the slow dynamics of a complex system to an underlying discrete Markovian process. This involves several steps, including space discretisation, time discretisation and dimensional reduction [2]. The analysed models in this thesis are constructed such that space discretisation is performed manually. All steps of the Markov State Model construction are presented on the minimal model introduced in the previous section 1.2.4.

1.3.1 Validate Markovian Process

Markov processes were introduced in section 1.2.2. An existing process is not bound to fulfill the Markov property so one has to check if it is met using the Chapman-Kolmogorov equation 1.29. We focus on Markov state models of NESS, so the time-dependence on Markovian transition probabilities is dropped. A constant time-step length of each Markovian jump called *lagtime* τ is assumed. The time-independent Chapman-Kolmogorov equation becomes

$$W(y_i|y_j, n\tau) = \int \mathrm{d}y_1 \int \mathrm{d}y_2... \int \mathrm{d}y_t W(y_1|y_i, \tau)W(y_2|y_1, \tau)...W(y_j|y_t, \tau) \quad (1.44)$$

for t time steps and the transition probabilities depend on the lagtime. The state space of an MSM is discretised so we denote $W(y_i|y_j, \tau) = p_{ij}(\tau)$ as the jump probability from state i to j within time τ. This quantity is estimated from simulation by choosing a lagtime τ and recording all jumps in a count matrix $c_{ij}(\tau)$. The transition matrix is estimated by

$$p_{ij}(\tau) = \frac{c_{ij}(\tau)}{\sum_k c_{ik}}. \quad (1.45)$$

In practice, it is cumbersome to check the Chapman-Kolmogorov equation for each element of the matrix $p_{ij}(\tau)$ so Prinz et al. [18] suggested to expand the irreducible transition probability matrix $\hat{p}(\tau)$ in its eigenvalue decomposition

$$\hat{p}\mathbf{\Phi}_k = \lambda_k \mathbf{\Phi}_k. \quad (1.46)$$

The time evolution of the probability distribution can then be described by

$$P(t) = \sum_k a_k \mathbf{\Phi}_k \exp\left(-\ln(\lambda_k)t\right), \quad (1.47)$$

where a_k are constants that are determined by the initial state of the system. If the dynamics of the system fulfil detailed balance (i.e. the system is in equilibrium), the eigenvalues and eigenvectors are real and one can construct a hierarchy starting with the slowest process $\lambda_0 = 1 > \lambda_1 > ... > 0$. Typically, the slow dynamical processes are of interested and the fast processes with small λ are cut off from the expansion. The characteristic timescales are identified from equation 1.47 by $t_i = \frac{-\tau}{\ln \lambda_i}$ and are calculated for different lagtimes. Figure 1.2 shows an example of the slowest timescales for varying lagtimes for 60 microstates. The region where $t_i < \tau$ is forbidden because an observed timescale cannot be smaller than the minimal time-resolution of the Markov process. The timescales reach a plateau for lagtime

greater than $0.002\,\mathcal{T}$, indicating that the Chapman-Kolmogorov equation 1.29 is valid in this region: An MSM at a timescale τ_0 is expected to describe the dynamics for $t > \tau_0$, so it should agree with an MSM parameterised at a lagtime $\tau_1 > \tau_0$ describing the same dynamics [28]. The timescale analysis is merely a tool to choose a lagtime for further analysis. A test of consistency of the eigenvectors is needed for full conformation of Markovianity. It is suggested to identify metastable states and compare detailed relaxation probabilities from the MSM to the error-analysed trajectory data. We start by defining the probability distribution of starting in a metastable state A by

$$\omega_i^A = \begin{cases} \frac{\pi_i}{\sum_{j \in A} \pi_j} & i \in A \\ 0 & i \notin A. \end{cases} \tag{1.48}$$

The probability of the measured trajectory data to remain in state A after time $m\tau$ is then

$$L_{\mathrm{tr}}(A, m\tau) = \sum_{i \in A} \omega_i^A \frac{\sum_{j \in A} c_{ij}(m\tau)}{\sum_j c_{ij}(m\tau)}, \tag{1.49}$$

with the corresponding error

$$\varepsilon(A, m\tau) = \sqrt{m \frac{L_{\mathrm{tr}}(A, m\tau) - (L_{\mathrm{tr}}(A, m\tau))^2}{\sum_{i \in A} \sum_j c_{ij}(m\tau)}}. \tag{1.50}$$

If the MSM dynamics follows the Markovian relation

$$\hat{p}(m\tau) \approx [\hat{p}(\tau)]^m, \tag{1.51}$$

the result of iterating the MSM

$$L_{\mathrm{MSM}}(A, \tau) = \sum_{i \in A} \left[(\omega_i^A)^\top [\hat{p}(\tau)]^m \right]_i \tag{1.52}$$

should lie within the errorbars of the trajectory analysis by $L_{\mathrm{tr}}(A, m\tau)$ with error $\varepsilon(A, m\tau)$. The method is shown on the 1D model in figure 1.3 for full conformation that the model defined by space discretisation and lagtime τ is Markovian for the slowest processes.

The eigenvalue decomposition allows detailed analysis of an MSM by isolating each process and showing detailed probability fluxes involved via the corresponding eigenvector. The described methods rely on the transition probability matrix being reversible, where the eigenvalue decomposition is real-valued. In NESS this condition
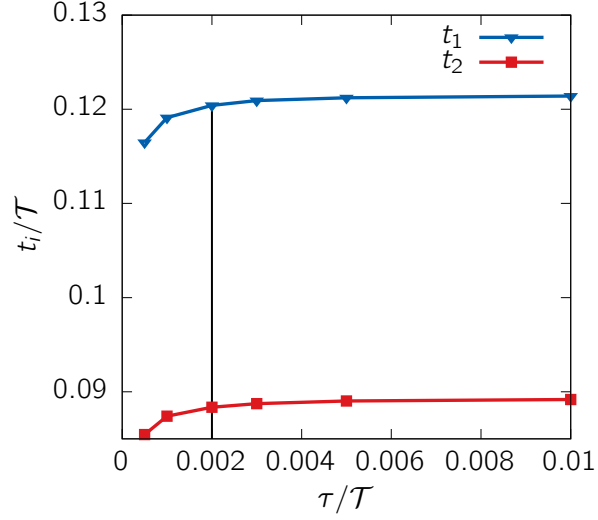
Figure 1.2: The slowest two timescales involved in the 1D system, depending on the lagtime τ of the MSM. A lagtime of $0.002\,\mathcal{T}$ is chosen for the following analysis.
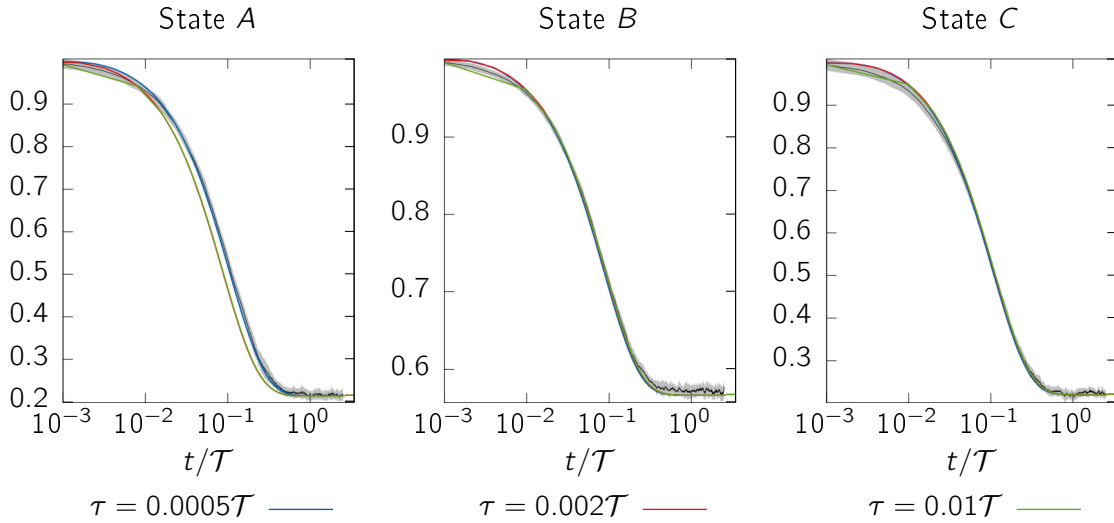

Figure 1.3: Chapman-Kolmogorov test for the metastable states A,B,C . The relaxation processes of the MSM at lagtime $\tau = 0.0005\,\mathcal{T}, 0.002\,\mathcal{T}, 0.01\,\mathcal{T}$ agree well within the error of trajectory analysis shown by the grey shadow.

is not met and the eigenvalues may become complex. A timescale separation with a hierarchy to delete fast processes with small real part of the eigenvalues is not possible [31]. A similarly powerful tool for analysis of NESS is not known yet, however the Schur-decomposition might be a good candidate for timescale separation [32]. In this thesis, a reference equilibrium system is used to choose a lagtime for all driven systems using the same potential surface. First-passage-time-distributions (FPTD) (see section 1.3.3) characterising the transition time distribution between two metastable states are used for analysis of the processes involved.

1.3.2   Identifying metastable states

Depending on the number and definition of microstates physical interpretation of MSMs can become difficult. The resulting fine-grained transition matrix does hold all information, but we are only interested in the previously unknown long-term dynamics. The already coarse-grained dynamics can be further coarse-grained to its relevant states. The aim is to combine dynamically well-connected microstates to a macrostate while keeping the timescale of the transitions constant. An often applied method in the field of molecular simulation is the PCCA+ [7], an expansion to the earlier introduced PCCA [6]. We will shortly explain the workflow of PCCA and point out the addition made by PCCA+.

By examination of the expansion of $P(t)$ in equation 1.47, one understands that the processes with $\lambda_i < 1$ vanish with time and only the stationary distribution Ψ_0 with $\lambda_0 = 1$ remains. The eigenvalue shows how fast the process annihilates in time, the corresponding eigenvector describes the corresponding probability flow. This is shown in figure 1.4 for the 1D toymodel. The metastable states are already identified by A, B and C. The probability flows from states where Ψ_i is positive to the negative region or vice versa, depending on what is needed to reach the steady state. The first eigenvector exchanges probability between metastable state A and C, the second between B and A, C equally, so the two flows cover all transition that one would intuitively expect from the system.

PCCA makes use of the eigenvector by a hierarchical approach, starting with the eigenvector corresponding to the slowest process. The positive region of Ψ_1 is defined as one metastable state and the negative region as another. It continues using the second slowest eigenvector and applies it in the same manner to the metastable state, where it is most active. It proceeds in such a way that the number of metastable
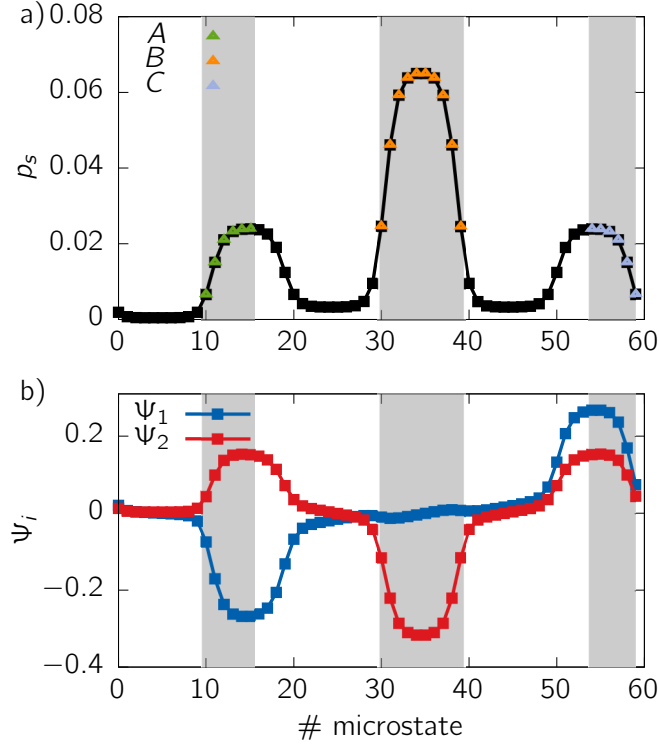
19

Figure 1.4: (a) Probability distribution for the equilibrium system with metastable states identified by PCCA+. (b) First and second eigenvector Ψ_i representing the probability flow during relaxation process.

states is defined by the number eigenvectors taken into account. This approach is error prone due to large inactive regions where $\Psi_i \approx 0$, that are randomly assigned to one or another metastable state. PCCA+ corrects for this mistake by assuming all eigenvectors at the same time and assigning a membership probability M for each microstate and each metastable state. A microstate between A and B would end up with membership of $M(A) \approx 50\%$, $M(B) \approx 50\%$ and $M(C) \approx 0\%$ due to its positioning as a transition state and should not belong to any metastable state. The states indicated in figure 1.4a each have membership $> 95\%$ for a state.

PCCA+ assigns metastable states that may not coincide with the maxima of a probability distribution as demonstrated for state A and C. It is focused on making the metastable states dynamically connected, crisp and well separated while keeping the transition timescales between the metastable states constant. To apply the algo-

rithm, the number of metastable states has to be chosen before running PCCA+. Another drawback is that it relies on the property that the eigenvalue expansion in equation 1.47 is real, which only holds for equilibrium system. Similar to the Markovianity check, the metastable states are determined once for the equilibrium system and assumed to be valid for the driven system too.

1.3.3   First-Passage-Time Distribution

First passage-times distributions (FPTD) are widely used to characterise processes in biology, chemistry and physics and are often associated with a free energy barrier a system has to overcome. The FPTD contains detailed transition information by collecting numerous realisations of a process. In experiment and simulation of rare events, the mean of the distribution is given due to limited observed process realisations [16]. Given an MSM with identified metastable states (see section 1.3.2) the FPTD between all metastable states can be calculated. The collection of initial states is denoted by I, of final states by F. For the purpose of calculating the FPTD [27] from I to F the MSM is modified such that all final microstates $f \in F$ become a sink, i.e. all jumps out of the metastable state have probability 0 and staying in the state has probability 1

$$
\begin{aligned}
\widetilde{p}_{fj} &= 0 & \forall j \neq f, & \quad \forall f \in F \\
\widetilde{p}_{ff} &= 1 & & \quad \forall f \in F.
\end{aligned}
\tag{1.53}
$$

An initial state is defined, where full probability is in one microstate $i \in I$ of the starting metastable state

$$
\begin{aligned}
\rho_i^{(0)} &= 1 & i \in I \\
\rho_j^{(0)} &= 0 & \forall j \neq i.
\end{aligned}
\tag{1.54}
$$

The probability distributions is iterated with the modified Markov model $\widetilde{p}$ until full probability is trapped in the sink. The superscript of (t) denotes the number of iterations performed.

$$
\rho^{(t+1)} = \widetilde{p}\rho^{(t)}.
\tag{1.55}
$$

The first-passage probability $p_{\text{FPT}}(t)$ at step t is then the probability flow in the sink

$$
p_{\text{FPT}}(i \rightarrow F, t) = \sum_{l \in F} \rho_l^{(t)} - \rho_l^{(t-1)}.
\tag{1.56}
$$

This calculation is repeated for all initial microstates $i \in I$. The full FPTD is then calculated by weighting each distribution with the weight w_i of a trajectory starting
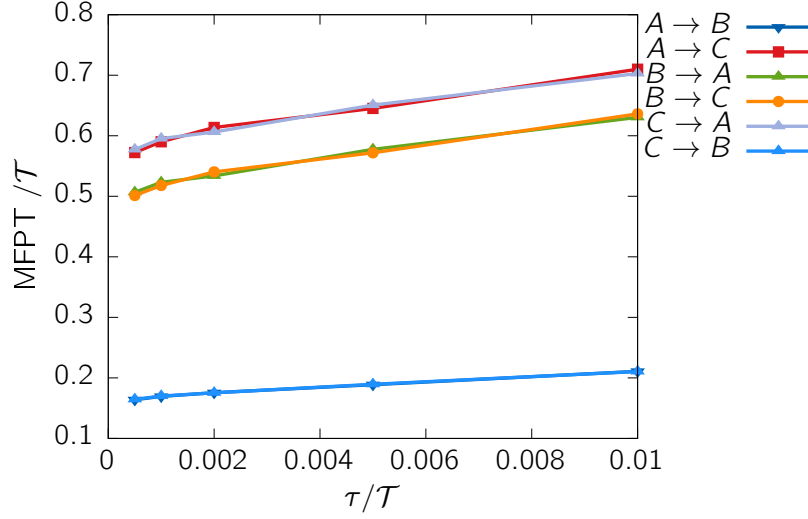
Figure 1.5: MFPT depending on the lagtime τ for all six observed processes. The model is designed such that groups of two processes overlap. The MFPT increases linearly with lagtime for all processes after a short period of faster growth.

in each initial microstate

$$w_i = \frac{\sum_{k \notin I} \pi_k p_{ki}}{\sum_{i \in I} \left(\sum_{k \notin I} \pi_k p_{ki} \right)}, \tag{1.57}$$

where $\boldsymbol{\pi}$ is the stationary probability distribution of a NESS. This expression adds up all probability flow to the initial state I by $\sum_{k \notin I} \pi_k p_{ki}$ and normalises the probability of flowing into macrostate I. Note that this expression does not use the modified MSM. The full FPT can then be expressed by

$$p_{\text{FPT}}(I \to F, t) = \sum_{i \in I} w_i p_{\text{FPT}}(i \to F, t). \tag{1.58}$$

Knowing the FPTD, all moments of the distribution can be calculated by

$$M_{I \to F}^{(n)} = \sum_t p_{\text{FPT}}(I \to F, t) t^n. \tag{1.59}$$
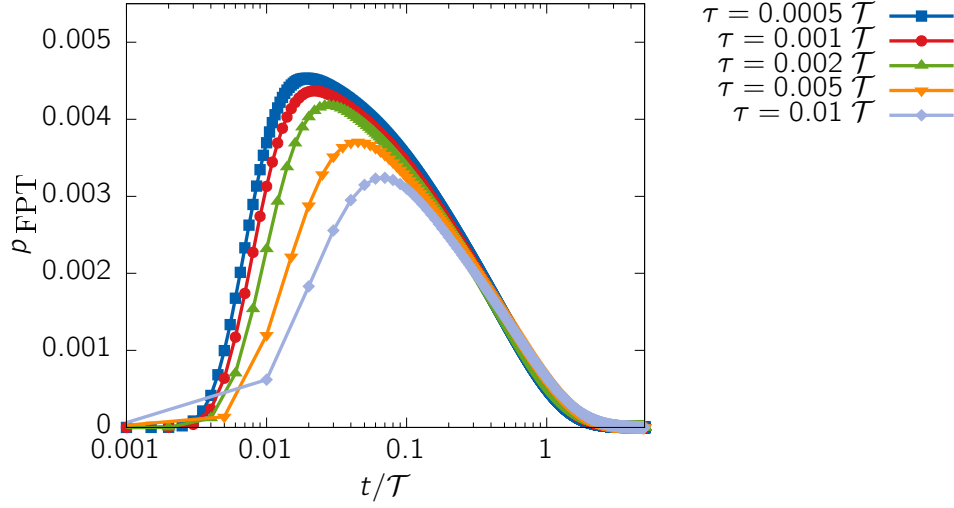
22

Figure 1.6: FPTD of process from metastable state B to A at different lagtimes. The distributions are normalised such that they can be compared despite different resolution. The distributions do not cover each other, as one would expect from a perfectly Markovian process.

In particular we will make use of the quantities

$$
\begin{aligned}
\mu_{I \to F} &= M_{I \to F}^{(1)} & \text{mean} \\
\sigma_{I \to F} &= \sqrt{M_{I \to F}^{(2)} - \mu_{I \to F}^2} & \text{standard deviation} \\
\kappa_{I \to F} &= \frac{M_{I \to F}^{(3)} - 3\mu_{I \to F}\sigma_{I \to F}^2 - \mu_{I \to F}^3}{\sigma_{I \to F}^3} & \text{standardised skewness,}
\end{aligned}
\tag{1.60}
$$

where the standardised skewness is defined by the expectation value of $\left(\frac{t-\mu}{\sigma}\right)^3$. The mean first-passage time (MFPT) was shown to be of special interested because it contains information about the transition rate between two states by the Hill relation [9]

$$
k_{I \to F} = \frac{1}{\mu_{I \to F}}. \tag{1.61}
$$

Similar to the timescale analysis and the Chapman-Kolmogorov test, the FPTD should not depend on the lagtime in the region of Markovianity. As shown in figure 1.5, this statement does not hold for the given system. The MFPT changes similar to the timescales in figure 1.2 up to a lagtime of $\approx 0.002\,\mathcal{T}$ but it continues

to grow linearly after this point. The reason for this non-Markovian behaviour is shown in the detailed FPTD in figure 1.6. The increasing lagtime summarises processes that are shorter than τ but still contribute to the overall MFPT as a single jump and details on the process are lost. The distribution is distorted to slower processes and the MFPT grows. To avoid this problem, the lagtimes should be chosen according to the criterions on Markovianity but as short as possible to keep details on the dynamics. The observed non-Markovianity is detected because the first-passage time is an even stronger test of Markovianity than the test performed in section 1.3.1. It analyses transitions between two metastable states and not just the exit from one. The presented exit test is used because it benefits from small error by making use of larger sample sizes. Many trajectories remain in the metastable state for a while. The error increases with time when the amount of trajectories remaining in the basin is small. The error for the FPTD is always large because it requires many incidents of the same transition time and a statistical analysis becomes cumbersome. Irrespective of the discussion of statistical applicability, the non-Markovianity of MFPT raises the question if it is a good quantity for comparison of simulation, MSM and experiment

Bibliography

[1] Valentin Blickle, Thomas Speck, Laurent Helden, Udo Seifert, and Clemens Bechinger. Thermodynamics of a colloidal particle in a time-dependent nonharmonic potential. *Physical review letters*, 96(7):070603, 2006.

[2] Gregory R Bowman, Vijay S Pande, and Frank Noé. *An introduction to Markov state models and their application to long timescale molecular simulation*, volume 797. Springer Science & Business Media, 2013.

[3] Gavin E Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. *Journal of Statistical Physics*, 90(5-6):1481–1487, 1998.

[4] Gavin E Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721, 1999.

[5] Bernard Derrida. Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07023, 2007.

[6] Peter Deuflhard, Wilhelm Huisinga, Alexander Fischer, and Christof Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. 1998.

[7] Peter Deuflhard and Marcus Weber. Robust perron cluster analysis in conformation dynamics. *Linear algebra and its applications*, 398:161–184, 2005.

[8] Peter Eastman. Introduction to statistical mechanics. CC BY-ND 4.0, 2015.

[9] Terrell L Hill. *Free energy transduction and biochemical cycle kinetics*. Courier Corporation, 2005.

[10] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A*, 31(3):1695, 1985.

[11] Wolfhard Janke. Monte Carlo simulations in statistical physics—from basic principles to advanced applications. In *Order, Disorder and Criticality: Advanced Problems of Phase Transition Theory Volume 3*, pages 93–166. World Scientific, 2013.

[12] Wilhelm Kutta. Beitrag zur naherungsweisen integration totaler differentialgleichungen. *Z. Math. Phys.*, 46:435–453, 1901.

[13] Ralf Labouvie, Bodhaditya Santra, Simon Heun, and Herwig Ott. Non-equilibrium steady-states in a driven-dissipative superfluid. *arXiv preprint arXiv:1507.05007*, 2015.

[14] Jan Liphardt, Sophie Dumont, Steven B Smith, Ignacio Tinoco, and Carlos Bustamante. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*, 296(5574):1832–1835, 2002.

[15] Xiao-guang Ma, Yun Su, Pik-Yin Lai, and Penger Tong. Colloidal dynamics over a tilted periodic potential: Forward and reverse transition probabilities and entropy production in a nonequilibrium steady state. *Physical Review E*, 96(1):012601, 2017.

[16] Nicholas F Polizzi, Michael J Therien, and David N Beratan. Mean first-passage times in biology. *Israel journal of chemistry*, 56(9-10):816–824, 2016.

[17] Yves Pomeau and Jarosław Piasecki. The Langevin equation. *Comptes Rendus Physique*, 18(9-10):570–582, 2017.

[18] Jan-Hendrik Prinz, Hao Wu, Marco Sarich, Bettina Keller, Martin Senne, Martin Held, John D Chodera, Christof Schütte, and Frank Noé. Markov models of molecular kinetics: Generation and validation. *The Journal of chemical physics*, 134(17):174105, 2011.

[19] Ch Schütte, Wilhelm Huisinga, and Peter 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.

[20] F. Schwabl. *Statistische Mechanik*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2006.

[21] Udo Seifert. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Physical review letters*, 95(4):040602, 2005.

[22] Udo Seifert. Stochastic thermodynamics: principles and perspectives. *The European Physical Journal B*, 64(3-4):423–431, 2008.

[23] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on progress in physics*, 75(12):126001, 2012.

[24] Ken Sekimoto. Langevin equation and thermodynamics. *Progress of Theoretical Physics Supplement*, 130:17–27, 1998.

[25] Thomas Soddemann, Burkhard Dünweg, and Kurt Kremer. Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Physical Review E*, 68(4):046702, 2003.

[26] Thomas Speck, Valentin Blickle, Clemens Bechinger, and Udo Seifert. Distribution of entropy production for a colloidal particle in a nonequilibrium steady state. *EPL (Europhysics Letters)*, 79(3):30002, 2007.

[27] Ernesto Suárez, Adam J Pratt, Lillian T Chong, and Daniel M Zuckerman. Estimating first-passage time distributions from weighted ensemble simulations and non-Markovian analyses. *Protein Science*, 25(1):67–78, 2016.

[28] William C Swope, Jed W Pitera, Frank Suits, Mike Pitman, Maria Eleftheriou, Blake G Fitch, Robert S Germain, Aleksandr Rayshubski, TJ Christopher Ward, Yuriy Zhestkov, et al. Describing protein folding kinetics by molecular dynamics simulations. 2. example applications to alanine dipeptide and a β-hairpin peptide. *The Journal of Physical Chemistry B*, 108(21):6582–6594, 2004.

[29] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

[30] Loup Verlet. Computer "experiments" on classical fluids. i. thermodynamical properties of Lennard-Jones molecules. *Physical review*, 159(1):98, 1967.

[31] Marcus Weber. Eigenvalues of non-reversible markov chains–a case study. 2017.

[32] Marcus Weber and Konstantin Fackeldey. G-PCCA: Spectral clustering for non-reversible markov chains. 2015.