GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Master's Thesis

Coarse graining of biochemical systems described by discrete stochastic dynamics

prepared by

David Seiferth

from Nienburg (Weser)

at the Institute for the Dynamics of Complex Systems

Thesis period:	28th January 2020 until 1st September 2020
Supervisor:	Prof. Dr. Stefan Klumpp
First referee:	Prof. Dr. Stefan Klumpp
Second referee:	Prof. Dr. Peter Sollich

Abstract

Many biological systems can be described by finite Markov models. A general method for simplifying master equations will be presented by merging two adjacent states. The method preserves the steady-state probability distribution and all steady-state fluxes disregarding the one between the merged states. Different levels of coarse-grained models of the underlying microscopic dynamics are obtained independently of the order of merging states. A criterion for the level of coarse graining or the resolution of the process will be proposed to find a sweet spot between the simplicity of a coarse-grained model and information loss compared to the original model.

Contents

1. Introduction

Finite-state stochastic Markov models play a major role in modelling biological systems with discrete states. To describe an N-particle system classically, one formulates the deterministic equations of motion. In principle, the equations of motion can be solved analytically. However, for high-dimensional systems, it may be cost intensive. Furthermore, the number of observables or rather the observed degrees of freedom are limited in an experiment. To take the unobserved degrees of freedom into account, one introduces stochastic variables. The description of such systems can be microscopic or macroscopic. Typically, the states of the system in the microscopic picture do not represent a complete classical or quantum-mechanical set of degrees of freedom. The states with observable degrees of freedom may have many internal degrees of freedom resulting from a coarse-graining procedure. The microscopic description can be related to macrostates of the system in terms of generalized thermodynamic forces and fluxes [1].

Coarse graining allows simplifying high dimensional dynamical systems by reducing the dimension of the state space. Moreover, coarse-grained models have computational and conceptual advantages compared to more detailed models. Recently, different coarse-graining approaches have been proposed, for example in [2–4]. Coarse graining goes along with information loss due to simplifying the original system. Most coarse-graining approaches define a method on how to simplify a system while retaining as much as possible key characteristics of the model. However, they often do not state how to identify microstates that can be faithfully coarse grained and which degree of coarse graining is optimal in some sense. Here, we present an iterative approach for finite-state Markov models of merging two adjacent states to obtain a hierarchy of models with different levels of coarse graining. Our approach does not impose any constraints on the network topology and hence is always applicable to merge states in systems governed by a master equation. By merging states, the dimensionality of the state space decreases and the system becomes simpler. But this goes along with information loss too. In our approach, the transition flux

between the merged states is lost and the dynamics of the system can be changed. Transition rates and steady-state probabilities are only changed locally and transition fluxes are retained. Since our coarse-graining procedure generates a hierarchy of models, we discuss a coarse-graining criterion which relates the information loss and simplicity. It can be used to find an optimal balance between information loss and simplicity and to find an optimised degree of coarse graining.

The thesis is organised as follows: After introducing the master equation and related quantities in section 2, we present our coarse-graining procedure and how transition rates in the coarse-grained system can be defined without changing the net numbers of transitions between any pair of states. We minimise the Kullback-Leibler divergence for trajectories to give another motivation for the choice of transition rates in the coarse-grained system. Our coarse-graining approach will be presented and discussed based on example systems in non-equilibrium steady states, first and foremost the molecular motor kinesin and its multi-cyclic kinetic diagram from Liepelt and Lipowsky [5]. First, we will coarse grain its kinetic diagram without changing its cycle topology in section 3.2. Secondly, we will discuss how different levels of coarse-grained models can be obtained by our coarse-graining approach and by reducing the number of cycles in the system in section 4.1. Apart from multi-cyclic models, uni-cyclic systems have been proposed to describe the molecular motor kinesin [6]. We will use our coarse-graining method to simplify a multi-cyclic kinesin model such that the coarse-grained model is uni-cyclic. We will show that our coarse-grained uni-cyclic model resembles the characteristics of other uni-cyclic models. Furthermore, we will present a criterion for the obtained different levels of coarse-grained models to find a sweet spot between simplicity of a coarse-grained model and information loss compared to the original model. Besides, the adaptation network of Chemotaxis of Escherichia coli (E. coli) bacteria will be discussed as an example of changed network topology in section 4.2.

2. Coarse-graining framework

2.1. Master equation and cycle decomposition

We consider discrete stochastic systems described by a master equation. The system has N discrete states and transition between states i and j occur with time-independent transition rates α_{ij}. At a certain time t, a state i is occupied with probability $p_i(t)$, which evolves according to the master equation

$$\frac{dp_i}{dt} = \sum_{j \neq i} (\alpha_{ji} p_j - \alpha_{ij} p_i) = \sum_{j \neq i} J_{ji}. \tag{2.1}$$

The master equation can be interpreted as balance of incoming ($\alpha_{ji} p_j$) and outgoing fluxes ($\alpha_{ij} p_i$). In a steady-state, the occupation probabilities are time-independent and hence the left-hand side of the equation is zero. In the following, we will focus on steady-state quantities unless stated otherwise. Steady-state probabilities can be calculated by matrix inversion of the master equation or in terms of a sum over spanning trees of a graph representing the states and transitions [7]. In equilibrium, there is no probability flux between any pair of states, a condition called detailed balance, as

$$\alpha_{ij} p_i = \alpha_{ji} p_j \tag{2.2}$$

for all pairs of states i and j. Here, we will consider systems in non-equilibrium steady-states. Hence, while the occupation probabilities are time-independent, $\partial_t p_i = 0$, the steady-state flux between a pair of states is in general not zero. The net flux from state i to j measures the net number of transitions per time and is given by

$$J_{ij} = \alpha_{ij} p_i - \alpha_{ji} p_j. \tag{2.3}$$

All nonzero fluxes are associated with the production of entropy, a feature that we will use below to quantify the information loss between a fine-grained and a coarse-grained description of the same system. The total entropy production P can be

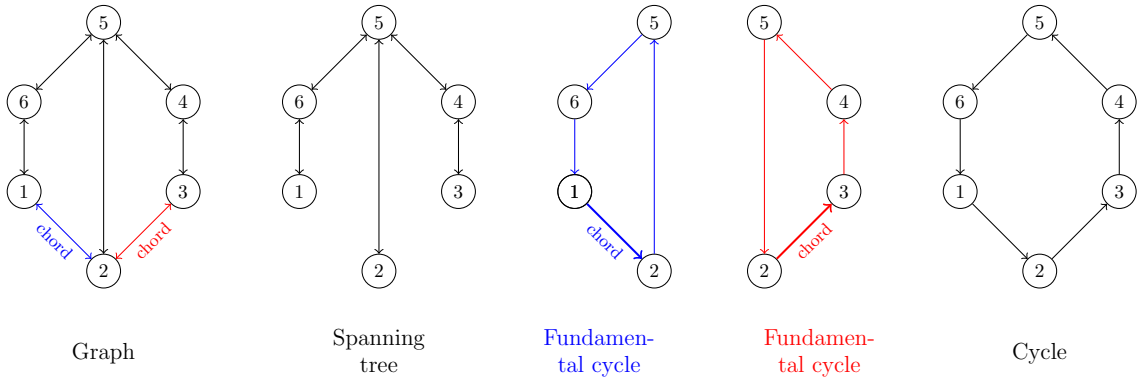


Figure 2.1.: A graph with one arbitrary chosen spanning tree. The edges missing in the spanning tree are called chords and define fundamental cycles. The blue coloured edge $1 \to 2$ is a chord and defines the blue cycle, while the red chord $3 \to 4$ defines the red cycle. The third cycle can be constructed by fundamental cycles [1]. The choice of a spanning tree of a graph is arbitrary. The reference orientation of all cycle is chosen to be counterclockwise.

expressed in terms of the net-transition fluxes J_{ij} and the corresponding transition affinities ΔS_{ij} [1],

$$P = \frac{1}{2} \sum_{i,j} J_{ij} \Delta S_{ij}. \tag{2.4}$$

The transition affinity

$$\Delta S_{ij} = \ln \left(\frac{\alpha_{ij} p_i}{\alpha_{ji} p_j} \right) \tag{2.5}$$

quantifies the change of entropy associated with a transition. The entropy production can be used to quantify the deviation from detailed balance for non-equilibrium systems [8].

For many purposes, it is useful to represent states and transitions as the nodes and edges of a graph. Doing so provides a constructive method for obtaining the steady state probabilities [1, 7]. Moreover, it allows us to decompose the dynamics in the network into a set of fundamental cycles and to express fluxes as well as the entropy production as sums over these fundamental cycles [1].

All nonzero fluxes in the steady state are associated with cycles, closed paths of nodes and edges in the graph. A large graph will typically contain multiple cycles, of which not all are independent of each other. In fact, all cycles can be constructed from a set of fundamental cycles, which in turn can be constructed from spanning trees. A spanning tree is a connected subgraph that contains all nodes, i.e. all states

of the system, but no cycles. It is obtained from the full graph by removing L edges. Adding back one edge to such a tree yields a single cycle. The added edge is called chord l and defines a corresponding fundamental cycle C_l. Figure 2.1 shows a 6-state network with three cycles of which two are fundamental. The choice of a spanning tree is arbitrary. Different spanning trees can result in different sets of fundamental cycles. The number of fundamental cycles is defined by the number L of chords of a network. In a network with E edges and N nodes, there are

$$L = E - N + 1 \tag{2.6}$$

chords [1]. Moreover, all transition fluxes can be expressed as a linear combination of chord fluxes [1]. Hence, a network with L chords has L independent transition fluxes and L fundamental cycles. There are two chords in figure 2.1 and hence two fundamental cycles. The entropy production can then be expressed in terms of quantities related to fundamental cycles or to the corresponding chords l as

$$P = \sum_l J_l \cdot \Delta S_{C_l}, \tag{2.7}$$

where the sum runs over a set of fundamental cycles. In this expression, J_l is the chord flux, the net-number of transitions between states connected by the chord l per unit time. ΔS_{C_l} is the affinity of the fundamental cycle C_l defined by chord l, that is the entropy change after one completion of the cycle. The cycle affinity can be written as

$$\Delta S_{C_l} = \sum_{(ij)\in C_l} \Delta S_{ij} = \ln \left(\prod_{(ij)\in C_l} \frac{\alpha_{ij}}{\alpha_{ji}} \right), \tag{2.8}$$

where both sum and product are evaluated over all edges (ij) within cycle C_l. We choose an arbitrary reference orientation for each fundamental cycle. In figure 2.1, the reference orientation is chosen to be counterclockwise.

2.2. Coarse-graining procedure

In this section, we present a procedure for coarse graining discrete stochastic systems. To simplify large networks, we merge two connected nodes m and n without losing important information or changing the dynamics of the system too much,

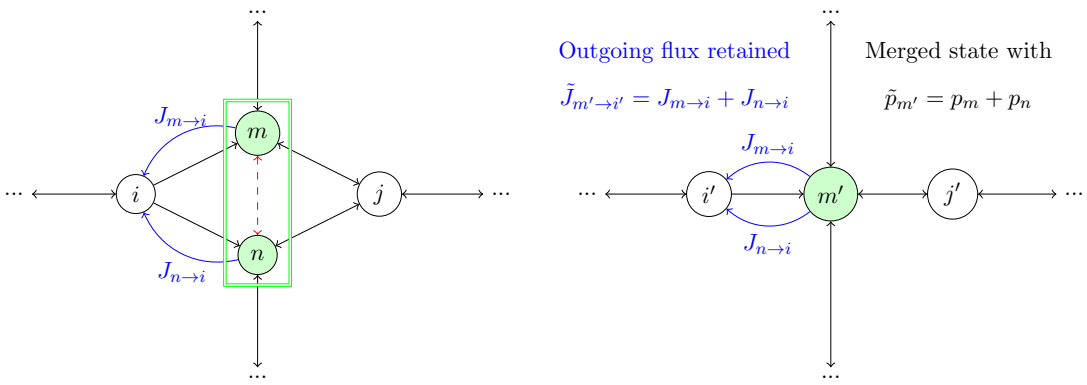


Figure 2.2.: State m and state n will be merged such that the fluxes between the merged state and the rest of the system are retained and the probabilities add up.

thus reducing the number of states by one. Further reduction can be achieved by iterating this procedure, as we will discuss below. The coarse-grained system shall still obey a master equation

$$\frac{d\tilde{p}_i}{dt} = \sum_{j\neq i} (\tilde{\alpha}_{ji}\tilde{p}_j - \tilde{\alpha}_{ij}\tilde{p}_i) \tag{2.9}$$

but with modified probabilities and transition rates. We adopt the following notation: As before, we denote the original probabilities and rates by p_i and α_{ij}. The probabilities of states in the coarse-grained system are denoted by $\tilde{p}_i$ and the transition rates by $\tilde{\alpha}_{ij}$. Note that in the following we derive constraints that refer only to steady-state probabilities. To conserve the probabilities, we request that the steady state probabilities of the two merged states add up, while all others are unaffected,

$$\tilde{p}_i = \begin{cases} p_i & \text{for } i \neq m, i \neq n \\ p_m + p_n & \text{for } i = m. \end{cases} \tag{2.10}$$

To determine the rates of the coarse-grained system, we impose the following additional requirements.

1. All transition rates that do not point to or from one of the two merged nodes remain unchanged, $\tilde{\alpha}_{ij} = \alpha_{ij}$ for $i, j \neq m, n$.

2. All one-way fluxes flowing from or to a merged node stay the same: outgoing (one-way) fluxes from the merged nodes n and m to the same node i are added, $\tilde{J}_{m\to i} = J_{m\to i} + J_{n\to i}$. Likewise, incoming (one-way) fluxes to the merged nodes

are added as well, $\tilde{J}_{i\to m} = J_{i\to m} + J_{i\to n}$.

Since the probabilities of nodes other than n and m stay the same, the latter condition directly implies that the incoming rate in the coarse-grained network is $\tilde{\alpha}_{im} = \alpha_{im}+\alpha_{in}$. The condition on outgoing fluxes determines the rates of transitions from the merged node to any other node.

Thus, all requirements stated above are fulfilled if the transition rates $\tilde{\alpha}_{ij}$ of the coarse-grained system are defined by

$$\begin{aligned} \tilde{\alpha}_{ij} &= \alpha_{ij} && \text{for } i,j \neq m,n \\ \tilde{\alpha}_{mj} &= \frac{\alpha_{mj}p_m + \alpha_{nj}p_n}{p_m + p_n} \\ \tilde{\alpha}_{im} &= \alpha_{im} + \alpha_{in}, \end{aligned} \tag{2.11}$$

where states m and n have been merged and states $i \neq m,n$ and $j \neq m,n$ are arbitrary states. Only transition rates pointing to or from the merged nodes are changed. The steady-state probabilities of the states that have not been merged stay the same and add up for the merged state. The transitions going into a merged state are a sum. The transition coming out of a merged state are a weighted sum of the transitions in the original system. All fluxes are retained except the one between the merged states.

Other quantities may be different in the coarse-grained system. These include the net-cycle fluxes, i.e. the number of cycle completions per time, the corresponding thermodynamic forces, and the energy dissipation.

The procedure of merging two adjacent states as described above can be iterated to further reduce the size of the system. Importantly, the result is independent of the order in which states are merged. We can differentiate two cases: On the one hand, we can merge first one pair of states and then another one, that has no direct transition to the first pair. Our approach incorporates only local changes and hence both coarse-graining iterations are independent. On the other hand, we can merge first a pair of states and then another one that is directly linked to the first pair. In that case, we can check by iterating equation (2.11) that merging states m and n with a third state k is associative.

As the described approach can be applied iteratively and the order of the steps does not affect the results, we can use it to obtain a hierarchy of different coarse-grained models of the same underlying microscopic dynamic. In general, we can define a mapping such that several states i in the original system are lumped together

in one state I. States in the original system are denoted by small letters and states in the coarse-grained system by capital letters. For a given mapping, the rates in the coarse-grained system are then given by

$$\tilde{\alpha}_{IJ} = \frac{\sum_{i\in I, j\in J} \alpha_{ij} p_i}{\sum_{i\in I} p_i}. \tag{2.12}$$

Here we have derived the probabilities and rates of the coarse-grained system from a set of plausible requirements imposed to make the steady-state properties of the simplified system similar to the ones of the original system. In section 2.5, we will give an alternative derivation and motivate the choice of transitions rates in the coarse-grained system by minimizing the Kullback-Leibler divergence for trajectories.

2.3. Coarse-grained energy landscape in equilibrium systems

We briefly consider the special case of canonical equilibrium systems. The probability distribution for systems in equilibrium is given by the Boltzmann distribution

$$p_i = \frac{1}{Z} e^{-\beta E_i}, \tag{2.13}$$

where Z is the partition sum, $\beta = 1/k_B T$ is the inverse of Boltzmann's constant and thermodynamic temperature and E_i the energy of state i. We compare the energy landscape of the original and coarse-grained system to make a consistency check of our approach. If we merge state m and state n in an equilibrium system, the probabilities add up:

$$\begin{aligned} \tilde{p}_{m'} &= p_m + p_n \\ &= \frac{1}{Z} \exp\left(\ln\left(e^{-\beta E_m} + e^{-\beta E_n}\right)\right) \\ &= \frac{1}{Z} e^{-\beta F_{m'}}. \end{aligned} \tag{2.14}$$

In the last expression, we have written the Boltzmann distribution of the coarse-grained state with a free energy instead of an energy. That this term indeed corresponds to a free energy $F_{m'} = \langle E_{mn} \rangle - TS$, given by the average energy $\langle E_{mn} \rangle$ of

the merged states and an internal entropy of the coarse-grained state can be seen as follows: Pulling the energy E_m out of the logarithm, we rewrite the free energy of the coarse-grained state as

$$\begin{aligned} F_{m'} &= E_m + k_B T \ln\left(\frac{e^{-\beta E_m}}{e^{-\beta E_m} + e^{-\beta E_n}}\right) \\ &= E_m + k_B T \ln\left(\frac{p_m}{p_m + p_n}\right). \end{aligned} \tag{2.15}$$

Likewise, pulling out E_n, we get

$$F_{m'} = E_n + k_B T \ln\left(\frac{p_n}{p_m + p_n}\right). \tag{2.16}$$

We can define $\bar{p}_m = p_m/(p_m + p_n)$ and $\bar{p}_n = p_n/(p_m + p_n)$ as re-scaled probabilities of the microscopic states within the coarse-grained state. Adding both equations with weights $\bar{p}_m$ and $\bar{p}_n$, respectively, yields

$$\begin{aligned} F_{m'} =& E_m \bar{p}_m + E_n \bar{p}_n + k_B T \left(\bar{p}_m \ln\left(\bar{p}_m\right) + \bar{p}_n \ln\left(\bar{p}_n\right)\right) \\ =& \langle E_{mn} \rangle - TS, \end{aligned} \tag{2.17}$$

where we expressed the free energy in terms of the average energy $\langle E_{mn} \rangle = E_m \bar{p}_m + E_n \bar{p}_n$ of the coarse-grained states and the internal entropy due to lumping two states together which is given by

$$S = -k_B \left[\bar{p}_m \ln\left(\bar{p}_m\right) + \bar{p}_n \ln\left(\bar{p}_n\right)\right]. \tag{2.18}$$

The latter arises due to the uncertainty, which micro-state in the merged state is occupied given that the system is in the coarse-grained state. The identification of the free energy of the merged states is consistent with the redefinition of the transition rates proposed in equation (2.11). The outgoing transition rates $\tilde{\alpha}_{mj}$ of a merged state m are weighted with with $\bar{p}_m$ and $\bar{p}_n$.

2.4. Comparison with other coarse-graining approaches

The coarse-graining approach we describe here is closely related to the so-called local equilibrium approximation [2]. Specifically, the expression we obtain for the rates in the coarse-grained system are equivalent to those obtained based on local equilibrium. However, there are differences in the justification and as a consequence in the applicability of the approach: The local equilibrium approximation makes use of time scale separation between degrees of freedom, see e.g. [2, 9]. A system in a non-equilibrium state often can be described as consisting of subsystems that are internally in equilibrium. The time evolution of each microstate within a macrostate is assumed to be the same [2]. Correspondingly, the local equilibrium approximation requires a time scale separation between rapid transitions between states that will be lumped together and slower transitions to the rest of the system. Typically, the states that are merged will be characterized by approximately equal energy levels. By contrast, our approach is not based on a time scale separation like the local equilibrium assumption. Instead, we imposed the requirements stated in equations (2.10) and (2.11) to preserve the net numbers of transitions between any pair of coarse-grained states. Our method thus retains the steady-state probability distribution during coarse-graining. It does not refer to equilibration or to the time scales associated with the dynamics. It can be applied generally to merge arbitrary neighboring states (including those that violate the requirements of the local equilibrium approximation) and will lead to a coarse-grained model that preserves the steady state distribution and fluxes. In principle, states with vastly different energy levels and states with a large flux between them can be merged as well. Likewise, transitions between the merged states do neither need to equilibrate rapidly to an equilibrium approximated by a Boltzmann distribution, nor do they need to relax rapidly into a non-equilibrium steady state. However, while our coarse-graining approach preserves the steady state of the system, it necessarily gives an approximation of the dynamics. The error of this approximation will be smallest if there is a time scale separation, i.e. if the requirements of the local equilibrium approximation are fulfilled. In that sense the requirements to preserve steady state quantities are similar to a local equilibrium approximation. However, they still allow the application of the approach in more general cases.

Several other coarse-graining approaches have been proposed recently [2–4, 10]

that differ in various aspects from each other and from the approach used here. In the following, we will briefly compare these approaches with our coarse-graining procedure. Table 2.1 summarises this comparison.

In [3], Pigolotti and Vulpiani presented an adiabatic approximation which eliminates rapidly evolving states from the system. All states with a mean dwell time [the inverse of the exit rate as given by equation (2.21)] below a certain threshold are removed and the rates of the remaining states are renormalized. The approximation is well suited for system with a time scale separation. In contrast to our approach, the calculation of a steady-state distribution is not needed. Hence, this approach is suitable e.g. for chemical networks with infinite state space, where calculating the steady state distribution is computationally challenging. However, the steady-state distribution is also not preserved in the coarse graining, as the probabilities of the states eliminated by coarse graining are redistributed globally and not only to the neighbouring states.

The approach by Altaner and Vollmer [4] is based on eliminating 'bridge states', states that can be eliminated in such a way that the cycle topology and the cycle affinities are preserved. In the coarse-graining step, the steady-state probability of the bridge state is distributed (not necessarily equally) to its two neighbouring states, such that the probabilities of non-neighbour states are unaffected (locality) and the systems' change in system entropy is preserved. This method is only applicable if the net flux along the bridge is nonzero. Hence, it is limited to non-equilibrium systems with nonzero currents.

In yet another approach [2], Hummer and Szabo proposed a coarse-graining method that leaves the occupancy-number correlation function $C_{ij}(t) = \langle \theta_i(t)\theta_j(0)\rangle$ unchanged. Here $\theta_i(t)$ is an indicator function that is one if state i is occupied at time t and zero otherwise. The calculation of the reduced matrix with the transition rates of the coarse-grained system requires matrix inversion and knowledge of the steady-state distribution as in our approach. However, in general, all transitions rates (not only the outgoing rates from the merged states) are changed. Hence, the method of Hummer and Szabo incorporates non-local changes of the transition matrix. The reduced matrix of the optimal Markov model is computed by Hummer and Szabo in the Markovian limit ($s \to 0$ in the Laplace space). They compare their reduced Markovian model with the local equilibrium approximation which corresponds to $s \to \infty$ in the Laplace space. If only two states are merged and hence all other macro states contain only one micro state, the local equilibrium approximation is

equivalent to the definition of transition rates in equation (2.11).

All coarse-graining procedures discussed so far are based on merging states. A different approach was proposed by Altaner et al. in [10]. Starting from the fact that edge fluxes can be represented as superpositions of cycle fluxes (see equation (3.16) for an example), a mapping is defined that transforms the original graph into a new one. The states in the new graph represent cycles of the original one. The transitions between the states representing cycles are defined in such way that the coarse-grained system fulfills detailed balance.

Our coarse-graining approach shares some features with these methods, but differs from them in others. For example, our approach requires the calculation of the steady-state distribution of the system, which can be computationally costly for large systems. The approach described by Pigolotti and Vulpiani in [3] is the only one of the methods presented here that does not require the steady-state distribution. In our approach, the state space is reduced via the merging of two arbitrary adjacent nodes. Adjacency is the only prerequisite, whereas the method of Altaner and Vollmer [4] is only applicable for bridge-states such that the cycle topology is preserved. The probability of the coarse-grained states is redistributed locally in our approach as well as in [4] whereas the methods of [2, 3] redistribute the probabilities non-locally. Likewise, the steady-state fluxes between states are not preserved in the adiabatic approximation with slow and fast states [3] but are retained in Altaner and Vollmer's [4] and in our approach. The cycle affinities (under the condition of preserved network topology) are preserved by the methods in [3, 4], but not by our approach. Due to the direct mapping in our coarse-graining approach, one can compare the original and the coarse-grained steady-state distribution. Furthermore, each node retains its meaning due to the direct mapping. A coarse-grained node contains at least two original nodes.

Table 2.1 summarizes and compares the coarse-graining approaches with respect to the properties discussed in this section. The comparison indicates that all coarse-graining approaches have their advantages and disadvantages and are better suited for some applications than others. Our approach is well suited for coarse-graining branches and for systems that can make use of the direct mapping between coarse-grained and original states. Branches are edges that are not part of any cycle and hence have zero steady-state transition fluxes. Besides, our approach is broadly applicable as it imposes no constraints on the network topology. Furthermore, due to the iterative nature of the approach, we can use it to obtain a hierarchy of models

Table 2.1.: Comparison of coarse-graining approaches.

	Adiabatic approximation with fast and slow states [3]	Fluctuation preserving coarse-grain method by Altaner and Vollmer in [4]	Approach by Hummer and Szabo in [2]	Our transition fluxes preserving coarse-grain approach
Assumption / required quantities	Outside steadstate regime valid too; calculation of steady state not needed	Steady state; requires non-zero flux along bridge that will be coarse-grained	Steady state	Steady state
Reduction of state space via	Bridge	Bridge	Lumping together N states in M lumped states	Merge 2 nodes
Redefinition of rates	Non-local changes of transition rates; affected by the dwell time in the bridge state	Proportion of one-way fluxes and proportion of probabilities define new rates from and to the node left and right of the bridge state	Non-local changes of transition rates; Matrix inversion of matrix product such that number correlation function is retained	Local changes: Outgoing flux from merged nodes preserved; all other rates retained
Cycle topology	Coarse graining of fundamental cycles possible, no restrictions	Preserved (requirement)	No restrictions	Coarse graining of branches and fundamental cycles possible
Local redistribution of probability	Non local redistribution	Only the steady-state prob. of the neighboring nodes changed	Probabilities within lumped states add up	Coarse-grained probability of merged node is the sum; all other prob. preserved
Affinities	Preserved	Preserved	Not preserved even if topology unchanged	Changed by the log of proportion of the one-way fluxes between merged states - see eq. (3.19)
Fluxes	Not preserved	Preserved	Not preserved	Preserved
Iterative	Commutes	Does not commute	Lumping order not important	Order of coarse-grain iterations not important (commutes)
Advantages	Well suited for infinite large networks (chemical networks)	Affinities preserved; can only coarse grain bridges with a non-zero, positive net-flux	Occupancy number correlation function retained	Coarse-graining of branches, direct mapping between coarse-grained and original states, merging two nodes is iterable and order is not important

that are coarse-grained to different degrees, a feature that we will use below.

2.5. Minimising the Kullback-Leibler divergence for trajectories

Another way to motivate the choice of transition rates in the coarse-grained system as given by equation (2.11) is by minimising the Kullback-Leibler divergence $\mathrm{KL}(q||\tilde{q})$ between probabilities q and $\tilde{q}$ of trajectories in the original system and in the coarse grained system, respectively [11]. The Kullback-Leibler divergence [12] is defined as

$$\mathrm{KL}(q||\tilde{q}) = \langle \ln \left(\frac{q}{\tilde{q}} \right) \rangle_q, \tag{2.19}$$

and provides a non-symmetric measure of the similarity of two probability distributions. For this derivation, we write the master equation in matrix notation

$$\frac{d\mathbf{p}}{dt} = \mathbf{W}\mathbf{p}, \tag{2.20}$$

where the matrix elements of $\mathbf{W}$ are given by

$$W_{ij} = \begin{cases} \alpha_{ji} & \text{for } i \neq j \\ -\sum_{k \neq i} \alpha_{ik} & \text{for } i = j. \end{cases} \tag{2.21}$$

The diagonal elements are called exit rates and ensure conservation of probability. The exit rate of state i is the sum of all outgoing transition rates. The master equation is solved by

$$\mathbf{p}(t) = e^{\mathbf{W}t}\mathbf{p}(0), \tag{2.22}$$

where $\mathbf{G}(t) = e^{\mathbf{W}t}$ denotes the propagator. To derive probabilities for trajectories, we discretize time and expanded the propagator for small time steps Δt: $\mathbf{G} = \mathbf{1} + \Delta t \mathbf{W}$ with

$$\begin{aligned} G_{ij} &= \Delta t \cdot W_{ij} \qquad (i \neq j) \\ G_{ii} &= 1 - \Delta t \sum_{j \neq i} W_{ji}. \end{aligned} \tag{2.23}$$

The continuous time dynamics described by the original master equation can be recovered in the limit of $\Delta t \to 0$ [13, 14]. A trajectory $\{i\} = (i_0, i_{\Delta t}, i_{2\Delta t}, ..., i_{t-\Delta t})$

is a sequence of states at different time steps. Its probability can be expressed in terms of the propagator as

$$q(\{i\}) = p_{i_0}(0) \prod_{\tau=0}^{t-\Delta t} G_{i_{\tau+\Delta t}, i_\tau}. \tag{2.24}$$

We now reduce the dimensionality of the state space by merging adjacent states. Instead of considering a single coarse-graining iteration as in section 2.2, we can examine a system after several coarse-graining iterations, as the order of those steps does not affect the result. After mapping states in the original system (denoted by small letters) to states in the coarse-grained system (denoted by capital letters), we have to define transition rates for the coarse-grained model $\tilde{W}_{IJ}$. We will show that the choice of transition rates in our coarse-graining approach minimises the Kullback-Leibler divergence $\mathrm{KL}(q(\{i\})||\tilde{q}(\{I\}))$, which compares the probability for trajectories in the original and the reduced state space. A trajectory $\{I\} = (I_0, I_{\Delta t}, I_{2\Delta t}, ..., I_{t-\Delta t})$ in the coarse-grained state space has probability

$$\tilde{q}(\{I\}) = \tilde{p}_{I_0}(0) \prod_{\tau=0}^{t-\Delta t} \tilde{G}_{I_{\tau+\Delta t}, I_\tau}, \tag{2.25}$$

where the short-time propagator for the reduced system has the matrix elements

$$\begin{aligned} \tilde{G}_{IJ} &= \Delta t\, \tilde{W}_{IJ} \qquad (I \neq J) \\ \tilde{G}_{II} &= 1 - \Delta t \sum_{J \neq I} \tilde{W}_{JI} \end{aligned} \tag{2.26}$$

in analogy to equation (2.23). In order to minimise the Kullback-Leibler divergence in (2.19), we have to consider the mean of the logarithm of the distribution $\tilde{q}$ in the coarse-grained system with respect to the distribution q in the original system:

$$\langle \ln \tilde{q} \rangle_q = \langle \ln \tilde{p}_{I_0}(0) \rangle_q + \sum_{\tau=0}^{t-\Delta t} \langle \ln(\tilde{G}_{I_{\tau+\Delta t}, I_\tau}) \rangle_q. \tag{2.27}$$

If we assume that the distribution p in the original system starts in its steady state, each time interval $(\tau, \tau + \Delta t)$ has the same average, such that

$$\langle \ln \tilde{q} \rangle_q = \sum_{i_0} p_{i_0} \ln \tilde{p}_{I_0}(0) + \frac{t}{\Delta t} \sum_{IJ} \sum_{i \in I, j \in J} G_{ij} p_j \ln(\tilde{G}_{IJ}) \tag{2.28}$$

the sum over the time steps in equation (2.27) can be simplified because $\langle \ln(\tilde{G}_{I_{\tau+\Delta t}, I_\tau}) \rangle_q$ becomes time independent in the steady-state and hence each summand in the sum over time contributes equally. The probabilities in a cluster I add up ($p_I = \sum_{i \in I} p_i$) and the cluster propagator can be defined as $G_{IJ} = p_J^{-1} \sum_{i \in I, j \in J} G_{ij} p_j$ such that equation 2.28 can be written as

$$\langle \ln \tilde{q} \rangle_q = \sum_{I_0} \tilde{p}_{I_0} \ln \tilde{p}_{I_0}(0) + \frac{t}{\Delta t} \sum_{IJ} G_{IJ} p_J \ln(\tilde{G}_{IJ}). \tag{2.29}$$

In the limit of long trajectories $t \to \infty$, the first term can be discarded. In the second term, we separate diagonal ($I = J$) and off-diagonal ($I \neq J$) terms:

$$\frac{1}{t} \langle \ln \tilde{q} \rangle_q = \sum_I G_{II} p_I \frac{\ln(\tilde{G}_{II})}{\Delta t} + \sum_{I \neq J} \frac{G_{IJ}}{\Delta t} p_J \left(\ln(\Delta t) + \ln \frac{\tilde{G}_{IJ}}{\Delta t} \right). \tag{2.30}$$

Going back to continuous time, i.e. for $\Delta t \to 0$, we can rewrite the factors in equation (2.30) such that

$$\begin{aligned} \lim_{\Delta t \to 0} G_{II} &= 1 \\ \lim_{\Delta t \to 0} \frac{G_{IJ}}{\Delta t} &= W_{IJ} \end{aligned} \tag{2.31}$$

and

$$\begin{aligned} \lim_{\Delta t \to 0} \frac{\tilde{G}_{IJ}}{\Delta t} &= \tilde{W}_{IJ} \\ \lim_{\Delta t \to 0} \frac{\ln(\tilde{G}_{II})}{\Delta t} &= \lim_{\Delta t \to 0} \frac{\ln(1 - \Delta t \sum_{J \neq I} \tilde{W}_{JI})}{\Delta t} = -\sum_{J \neq I} \tilde{W}_{JI} \end{aligned} \tag{2.32}$$

where W_{IJ} defines the clustered rates of the original model and $\tilde{W}_{IJ}$ the approximating rates in the coarse-grained model. The term with $\ln(\Delta t)$ cancels with the same term in $\langle \ln q \rangle_q$ in the Kullback-Leibler divergence $\mathrm{KL}(q||\tilde{q}) = \langle \ln q \rangle_q - \langle \ln \tilde{q} \rangle_q$. After performing the limit, equation (2.30) can be written as

$$\begin{aligned} \lim_{\Delta t \to 0} \frac{1}{t} \langle \ln \tilde{q} \rangle_q &= -\sum_I p_I \sum_{J \neq I} \tilde{W}_{JI} + \sum_{I \neq J} W_{IJ} p_J \ln(\tilde{W}_{IJ}) + c \\ &= \sum_{I \neq J} (W_{IJ} p_J \ln(\tilde{W}_{IJ}) - \tilde{W}_{IJ} p_J) + c \end{aligned} \tag{2.33}$$

with a constant c that is independent of of $\tilde{W}_{IJ}$. The minimisation of $\mathrm{KL}(q||\tilde{q})$ is equivalent to maximising its second term $\langle \ln \tilde{q} \rangle_q$. The derivative with respect to $\tilde{W}_{IJ}$

of the second term $\langle \ln \tilde{q} \rangle_q$ of the Kullback-Leibler divergence is

$$\frac{d}{d\tilde{W}_{IJ}} \frac{1}{t} \langle \ln \tilde{q} \rangle_q = p_J \left(\frac{W_{IJ}}{\tilde{W}_{IJ}} - 1 \right) = 0. \tag{2.34}$$

Hence, $\tilde{W}_{IJ} = W_{IJ}$ minimises the Kullback-Leibler divergence $\mathrm{KL}(q||\tilde{q})$ and yields the same choice for transition rates for the coarse-grained system

$$\tilde{W}_{IJ} = W_{IJ} = \frac{\sum_{i \in I, j \in J} G_{ij} p_j}{\sum_{j \in J} p_j} \tag{2.35}$$

as in equation (2.11). Summarising, then, we have shown in this section that minimising the Kullback-Leibler divergence for trajectories yields the same choice of transition rates as proposed in equation (2.11).

3. Coarse graining with preserved cycle topology

In case the detailed balance condition is not fulfilled, there are non-zero net fluxes in the system and the steady state does not obey the Boltzmann distribution anymore. To examine the effect of the coarse-graining procedure on non-equilibrium steady states (NESS) of uni-cyclic systems in section 3.1, we differentiate between two cases. The first case concerns coarse graining of branches. Herein, the net-transition flux in a branch is zero in a steady state. The thermodynamic force driving a cycle in a non-equilibrium system remains unchanged in the coarse-grained system. This particular setting will be analysed in section 3.1.1. Alternatively, the net flux of a transition within a cycle in NESS systems is non-zero. This means, if two states within a cycle are merged, the flux between them will be lost. The effects of coarse graining within a cycle will be described section 3.1.3. Notably, the thermodynamic force driving a cycle, which is not in equilibrium, does not necessarily remain unchanged in the coarse-grained system. In section 3.2, we will discuss how a multi-cyclic network can be coarse grained without changing its cycle topology.

3.1. Coarse graining of uni-cyclic networks

The Boltzmann distribution for equilibrium systems in equation (2.13) allowed us to assign energy levels E_i to states i that are occupied with probability p_i. To assign energy levels to states in a non-equilibrium system, we relate the ratio of forward and backward transition to energy levels and work done by a force driving the transition, i.e. in conjunction with

$$\frac{\alpha_{ij}}{\alpha_{ji}} = e^{\beta(E_i - E_j + u_{ij})},$$

where E_i is an energy level and u_{ij} is the work done by an external force.

3.1.1. Coarse-grained branches

Adding one edge and one node to a three-state cycle yields a cycle with one branch as depicted in figure 3.1. In this subsection, the steady state of this system will be calculated for a non-equilibrium example and the branch will be coarse-grained. All rates α_{ij} remain constant except for $\tilde{\alpha}_{21}$ and $\tilde{\alpha}_{20}$ which represent the outgoing rates from the merged state 2 depicted in red in figure 3.1. The outgoing rates of the new state 2 are characterised by the constraint that the outgoing fluxes have to be retained as defined in equation (2.11):

$$\begin{aligned} \alpha_{21} \cdot p_2 &= \tilde{\alpha}_{21} \cdot \tilde{p}_2 \\ \alpha_{20} \cdot p_2 &= \tilde{\alpha}_{20} \cdot \tilde{p}_2. \end{aligned}$$

Hence, the outgoing rates of the new state 2 are given by

$$\begin{aligned} \tilde{\alpha}_{21} &= \alpha_{21} \frac{p_2}{\tilde{p}_2} = \frac{\alpha_{21}\alpha_{32}}{\alpha_{32} + \alpha_{23}} \\ \tilde{\alpha}_{20} &= \alpha_{20} \frac{p_2}{\tilde{p}_2} = \frac{\alpha_{20}\alpha_{32}}{\alpha_{32} + \alpha_{23}}. \end{aligned} \tag{3.1}$$

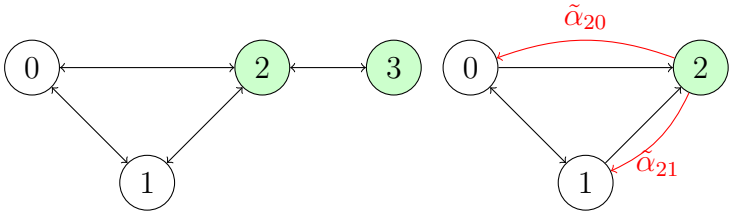


Figure 3.1.: The green-coloured states in the branch are merged. The transition rates in the coarse-grained system are defined according to equation (2.11). The outgoing rates from the merged state $\tilde{\alpha}_{20}$ and $\tilde{\alpha}_{21}$ are explicitly calculated in equation (3.1).

Herein, the thermodynamic force driving the cycle is unaffected by the coarse-grained branch as the ratio of the transition rates in the forward and backward direction of the cycle is not changed. Since there is only one cycle in the network, the net-edge flux is the same as the net-cycle flux (net number of cycle completions per time), which is unaffected by the coarse-graining procedure as well. In analogy to the effective energy defined in section 2.3 for a system in equilibrium, we seek to assign an effective energy for the coarse-grained state 2 which represents in the original network states 2 and 3. State i with $i \in \{0, 1, 2, 3\}$ has potential energy E_i and work is done by an external force during the transition $0 \to 1$ such that the

ratio of transition rates are

$$\begin{aligned} \frac{\alpha_{01}}{\alpha_{10}} &= e^{\beta(E_0 - E_1 + u_{01})} & \frac{\alpha_{12}}{\alpha_{21}} &= e^{\beta(E_1 - E_2)} \\ \frac{\alpha_{20}}{\alpha_{02}} &= e^{\beta(E_2 - E_0)} & \frac{\alpha_{23}}{\alpha_{32}} &= e^{\beta(E_2 - E_3)}. \end{aligned}$$

By construction, this yields a thermodynamic force

$$\Delta S_C = \ln\left(\frac{\alpha_{01}\alpha_{12}\alpha_{20}}{\alpha_{10}\alpha_{02}\alpha_{21}}\right) = u_{01}$$

driving the cycle. To give an example, we consider the network in figure 3.1 with the following rates $\alpha_{ij} = (A)_{ij}$ given by

$$A = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{3.2}$$

where the row i belong to the rates pointing away from state i. The columns denote the rates pointing to state j. By equation (3.1) the coarse-grained network has the following rates:

$$\tilde{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 3 \\ 1/3 & 1/3 & 0 \end{pmatrix}.$$

The dimensionality of $\tilde{A}$ is reduced by one compared to A since we merged state 2 and 3. The first and second row remain constant because state 0 and state 1 are not affected by the coarse graining. The corresponding steady-state probabilities are

$$\begin{aligned} \vec{p} &= \begin{pmatrix} 1/8 & 1/8 & 1/4 & 1/2 \end{pmatrix} \text{ and} \\ \tilde{\vec{p}} &= \begin{pmatrix} 1/8 & 1/8 & 3/4 \end{pmatrix}. \end{aligned}$$

In particular, the steady-state probabilities for state 0 and state 1 remain constant. The net-flux (counterclockwise) is $J = 1/8$ and is invariant in the coarse-grained system by construction as requested in section 2.2. The ratio of fluxes in the counterclockwise and clockwise direction is related to the thermodynamic force (or the

affinity of the cycle C)

$$\Delta S_C = \ln\left(\frac{\alpha_{01}\alpha_{12}\alpha_{20}}{\alpha_{10}\alpha_{02}\alpha_{21}}\right) = \ln 6 = \Delta\tilde{S}_C,$$

which is unaffected by the coarse-graining procedure as well. In the next section, the dwell time distribution for the branch and the state connecting the cycle and the branch will be calculated. Furthermore, it will be investigated how coarse graining affects the dwell time distribution.

3.1.2. Dwell time distribution in the coarse-grained state

The time, in which a state is occupied, is exponentially distributed with exit rate

$$\lambda_i = \sum_{j=0, j\neq i}^{N-1} \alpha_{ij}, \tag{3.3}$$

where N is the number of states in the network. The exit rate is the sum of all elements of a row in a matrix, i.e. the exit rate of state i is the sum of all transition rates pointing away from state i. The probability distribution function $\rho(t)$ for the waiting time is given by

$$\rho(t) = \lambda_i \exp(-\lambda_i t). \tag{3.4}$$

The mean time in state i before a transition can be expressed as

$$\langle t_i \rangle = \int_0^\infty t\lambda_i \exp(-\lambda_i t)dt = \frac{1}{\lambda_i}. \tag{3.5}$$

For the coarse-grained network the exit rate is determined by

$$\tilde{\lambda}_i = \sum_{j=0, j\neq i}^{N-2} \tilde{\alpha}_{ij} = \frac{p_i}{\tilde{p}_i} \sum_{j=0, j\neq i}^{N-2} \alpha_{ij} < \lambda_i. \tag{3.6}$$

The coarse-grained state has one less outgoing transition because it has no transition to the branch anymore. Furthermore, the probability for the merged state is larger than the one for the fine-grained state ($p_i < \tilde{p}_i$). Therefore, the factor $p_i/\tilde{p}_i < 1$ in equation (3.6) is smaller than one. Taking this into account, the exit rate $\tilde{\lambda}_i$ for a merged state is smaller than for the original state. Consequently, the average

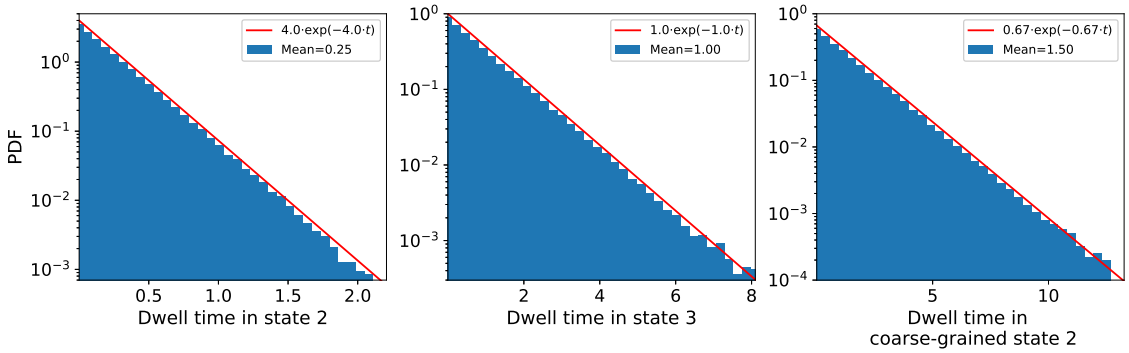


Figure 3.2.: Distribution of the dwell time of several states of the system depicted in figure 3.1. The transition rates are defined in equation (3.2). The blue histograms are generated from a Gillespie simulation. The red line represents the analytically expected exponential distribution. The mean time in state i is inversely proportional to the sum of all outgoing rates from state i (the exit rate), as given by equation (3.5).

waiting time or also called dwell time in the merged state is higher than in the state in the original network.

$$\langle \tilde{t}_i \rangle > \langle t_i \rangle. \tag{3.7}$$

Figure 3.2 shows the dwell time distributions for states 2 and 3 in the fine-grained model and the dwell time distribution for the coarse-grained state 2. With the parametrisation from equation (3.2), the mean waiting time in state 2 in the fine-grained model is given by

$$\langle t_2 \rangle = \frac{1}{\lambda_2} = \frac{1}{\alpha_{20} + \alpha_{21} + \alpha_{23}} = \frac{1}{4}.$$

The mean dwell time in state 3 can be expressed as

$$\langle t_3 \rangle = \frac{1}{\lambda_3} = \frac{1}{\alpha_{32}} = 1.$$

In the coarse-grained system, state 2 and 3 are merged. The merged state has a mean waiting time

$$\langle \tilde{t}_2 \rangle = \frac{1}{\tilde{\lambda}_2} = \frac{1}{\tilde{\alpha}_{20} + \tilde{\alpha}_{21}} = \frac{\alpha_{23} + \alpha_{32}}{\alpha_{32}(\alpha_{20} + \alpha_{21})} = \frac{3}{2}. \tag{3.8}$$

The distribution of the dwell time in the branch is calculated explicitly in the appendix in section A and plotted in figures 3.3 and A.1. We can distinguish two time scales in the distributions. Figures 3.3 and A.1 both show an exponential distribution of the dwell time for long times. However for short dwell times, the dwell time distribution in the branch cannot be described by the same exponential distribution. To calculate the dwell time distribution analytically by using convolutions in the appendix, we consider a special case where the exit rates $\lambda_2 = \lambda_3$ are equal. Let us emphasize that our characterisation of the n-fold convolution requires $\lambda_2 = \lambda_3$ and does not easily generalise to the general setting of $\lambda_2 \neq \lambda_3$. If the sum of outgoing rates from state 2 and 3 are the same ($\lambda_2 = \lambda_3$) the waiting time distribution is described by equation (A.5) in the appendix

$$\rho(t) = \frac{\alpha_{20} + \alpha_{21}}{2} \left(e^{-(\lambda - \sqrt{\alpha_{23}\lambda})t} + e^{-(\lambda + \sqrt{\alpha_{23}\lambda})t} \right).$$

The two time scales can be read off immediately. The first term decaying with $\lambda - \sqrt{\alpha_{23}\lambda}$ is dominant for large t because the second term decays faster with $\lambda + \sqrt{\alpha_{23}\lambda} > \lambda - \sqrt{\alpha_{23}\lambda}$. For large t, the waiting time distribution is exponentially distributed with $\rho(t) \sim e^{-(\lambda - \sqrt{\alpha_{23}\lambda})t}$. If the outgoing rates of state 2 and 3 are not the same, the dwell time distribution is more coplicated to calculate and is described by equation (A.7) in the appendix

$$\rho(t) = \frac{\alpha_{20} + \alpha_{21}}{\lambda_2} \left(\lambda_2 e^{-\lambda_2 t} + \sum_{n=1}^{\infty} \rho_{conv}(n, \lambda_2, \lambda_3, t) \left(\frac{\alpha_{23}}{\lambda_2} \right)^n \right).$$

For very small t, the first term $(\alpha_{20} + \alpha_{21}) \cdot e^{-\lambda_2}$ is dominant. In the middle plot of figure 3.3, the first term is depicted in blue and describes the distribution well for very small t. The exponential decay rate for large time scales can be obtained via a fit which is coloured in red in figure 3.3. For short dwell times, the dwell time distributions for the branch and the coarse-grained state exhibit differences. For long dwell times, the differences between the distribution decrease. Our coarse-graining approach is not designed to preserve dwell time distributions. Coarse graining branches changes the dynamics of the system. The non-exponential dwell time distribution for branches is not retained by our coarse-graining approach. However, the mean dwell time for the coarse-grained state is the same as for the branch in the original system. Explicit calculations of the mean dwell time of the distributions can be found in equations (A.6) and (A.9). Here, we briefly discussed how

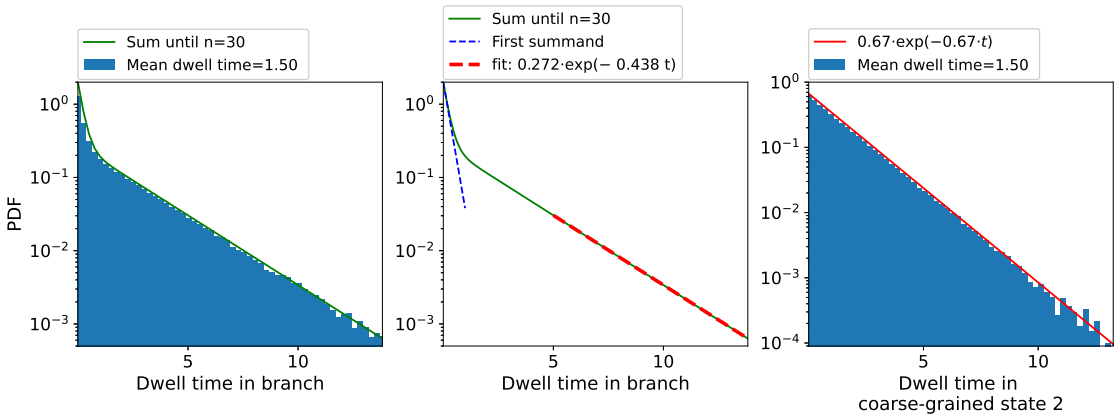


Figure 3.3.: Distribution of the time spent in the branch in the fine-grained system and in the merged state in the coarse-grained system. The blue histograms are generated from a Gillespie simulation. Each histogram contains $2.5 \cdot 10^5$ data points with 100 bins. Both histograms describe states 2 and 3 as one state. The first plot is a network without new rates. The simulation differs between state 2 and 3 but state 2 and 3 are analyzed as the same state, i.e. state 3 is an internal state of state 2. The green line in the first and second plot refers to equation (A.7). The sum is only evaluated until $n = 30$. For large t the distribution is exponential with parameters which can be fitted. The transition rates are the same as in equation (3.2). The third plot shows the exponentially distributed waiting time in state 2 in the coarse-grained system, where states 2 and 3 are merged and the outgoing rates of the merged state 2 are given by $\tilde{\alpha}_{20}$ and $\tilde{\alpha}_{21}$ in equation (3.1). The red line represents the analytically expected result of the exponential dwell time distribution.

coarse graining affects the dwell time in branches. The effect of coarse graining on dwell time distributions in multi-cyclic systems is discussed in section B.4 in the supplementary material.

3.1.3. Coarse graining within cycles

After discussing coarse-grained branches, we investigate how merging states within cycles changes the system. The system depicted in figure 3.4 has one cycle and four states. The system can be interpreted as a generic model for enzymatic catalysis, where the states correspond to different bound molecules (e.g. ATP, ADP and P). The model for the molecular motor kinesin introduced by Fisher and Kolomeisky [6] has the same topology as the 4-state cycle in figure 3.4. In this generic 4-state

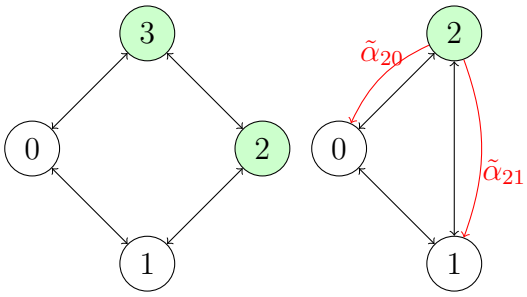


Figure 3.4.: The green-coloured states in the uni-cyclic 4-state network are merged in (a). The coarse-grained network in (b) is still uni-cyclic. The red transition rates are defined in equation (3.11) analogous to equation (2.11) to conserve the outgoing flux from state 3 to state 0 and from state 2 to state 1.

cycle, the forward and backward transitions obey

$$\frac{\alpha_{ij}}{\alpha_{ji}} = e^{\beta(E_i - E_j + u_{ij})}. \tag{3.9}$$

Each state i has a potential energy E_i. An external force does work $u_{ij} = -u_{ji}$ during a transition from i to j. The thermodynamic force driving the cycle is given by

$$\begin{aligned} \Delta S_C &= \ln\left(\frac{\alpha_{01}\alpha_{12}\alpha_{23}\alpha_{30}}{\alpha_{10}\alpha_{21}\alpha_{32}\alpha_{03}}\right) \\ &= u_{01} + u_{12} + u_{23} + u_{30}. \end{aligned} \tag{3.10}$$

Figure 3.4 shows how a cycle consisting of four states is coarse grained to a 3-state cycle. The merged node is not a branch and is connected via two edges with the rest of the system. The steady-state probabilities p_0 and p_1 are unaffected by the coarse graining if one defines the outgoing rates of the new state 2 according to the procedure from equation (2.11) as

$$\begin{aligned} \tilde{\alpha}_{21} &= \alpha_{21}\frac{p_2}{p_2 + p_3} \\ \tilde{\alpha}_{20} &= \alpha_{30}\frac{p_3}{p_2 + p_3}. \end{aligned} \tag{3.11}$$

The transition rate from state 0 to state 3 gets rewired such that $\tilde{\alpha}_{02} = \alpha_{03}$. All other transition rates remain unchanged. The thermodynamic force driving the

coarse-grained cycle is given by

$$\Delta\tilde{S}_C = \ln\left(\frac{\tilde{\alpha}_{01}\tilde{\alpha}_{12}\tilde{\alpha}_{20}}{\tilde{\alpha}_{10}\tilde{\alpha}_{21}\tilde{\alpha}_{02}}\right). \tag{3.12}$$

Note that in general, the force driving the cycle does not remain unchanged, even if no work $u_{23} = 0$ is done in the neglected transition $2 \to 3$ between merged states. The net-cycle flux (i.e. the numbers of cycle completions per time) is conserved because the coarse-grained network has by construction the same net transition fluxes. In a uni-cyclic network, net-edge and net-cycle flux are identical. The thermodynamic force in the coarse-grained system can be expressed in terms of the one in the original system as

$$\Delta\tilde{S}_C = \ln\left(\frac{\tilde{\alpha}_{01}\tilde{\alpha}_{12}\tilde{\alpha}_{20}}{\tilde{\alpha}_{10}\tilde{\alpha}_{21}\tilde{\alpha}_{02}}\right) = \Delta S_C - \ln\left(\frac{\alpha_{23}p_2}{\alpha_{32}p_3}\right) \tag{3.13}$$

where both thermodynamic forces are equal, if $\alpha_{23}p_2 = \alpha_{32}p_3$. This is equivalent to a zero net-flux between state 2 and 3 or detailed balance. The difference in entropy production between the four-node and the three-node cycle is given by

$$\Delta P = P - \tilde{P} = J_{23}\ln\left(\frac{\alpha_{23}p_2}{\alpha_{32}p_3}\right). \tag{3.14}$$

The contribution of the merged edge $(2, 3)$ is equal to the difference in entropy production. The affinities associated with the edges of the merged node 2 can be expressed in terms of the original system as

$$\begin{aligned}\Delta\tilde{S}_{20} &= \ln\frac{\tilde{\alpha}_{20}\tilde{p}_2}{\alpha_{02}\tilde{p}_0} = \Delta S_{30}\\ \Delta\tilde{S}_{12} &= \ln\frac{\tilde{\alpha}_{12}\tilde{p}_1}{\alpha_{21}\tilde{p}_2} = \Delta S_{12}.\end{aligned} \tag{3.15}$$

3.2. Coarse graining multi-cyclic networks - the molecular motor kinesin

In this and the next section, we will discuss the coarse graining of the kinetic diagram of the molecular motor kinesin as an example for simplifying a system in a non-equilibrium steady-state. We will first coarse grain the network without changing

its cycle topology and postpone coarse-graining steps to remove cycles to the next section.

3.2.1. Cycle decomposition of the kinetic diagram for the kinesin motor

Kinesin is a motor protein with two heads which can carry cargo, moves along microtubule filaments and is powered by the hydrolysis of adenosine triphosphate (ATP) [15, 16]. Kinesin can be described by the 6-states network [5] that we have already used as an example above, depicted in figure 3.5. The six states correspond to different chemical states of the two motor heads (ATP- or ADP-bound or free). Transition $2 \to 5$ corresponds to a mechanical forward step, the backward transition $5 \to 2$ to a backward step. Alternative mechanical stepping transitions have been discussed [17], but will not be considered here. The other transitions describe chemical changes in the motor heads, i.e. binding or release of ATP, or ADP and the hydrolysis of ATP to ADP in the front head and rear head (transitions $6 \to 1$ and $3 \to 4$, respectively) including the release of phosphate.

The network can be decomposed into cycles as described in the previous section. The blue-coloured cycle in figure 2.1 ($C_F = \{1, 2, 5, 6, 1\}$) includes a mechanical forward transition and an ATP hydrolysis transition and is hence called forward cycle C_F. It corresponds to the normal working cycle of the kinesin motor under low loads. The red coloured cycle $C_B = \{2, 3, 4, 5, 2\}$ incorporates a mechanical backward transition and an ATP hydrolysis transition and is called backward cycle C_B. The third cycle, which is not fundamental in the decomposition in figure 2.1, contains no mechanical transitions but two ATP hydrolysis transitions and hence is a dissipative or futile cycle C_D. The network has two independent transition fluxes. With the choice of two chords in the cycle decomposition depicted in figure 2.1, all transition fluxes can be expressed in terms of the chord fluxes J_{12} and J_{23} as

$$\begin{aligned} J_{12} &= J_{56} = J_{61} = J_F + J_D \\ J_{23} &= J_{34} = J_{45} = J_B + J_D \\ J_{25} &= J_{12} - J_{23} = J_F - J_B. \end{aligned} \tag{3.16}$$

For the sake of completeness, the relations between transition and cycle fluxes are listed as well. The cycle fluxes J_C are defined according to Hill's convention [7] and can be interpreted as the number of completions of cycle C per time. Cycle fluxes

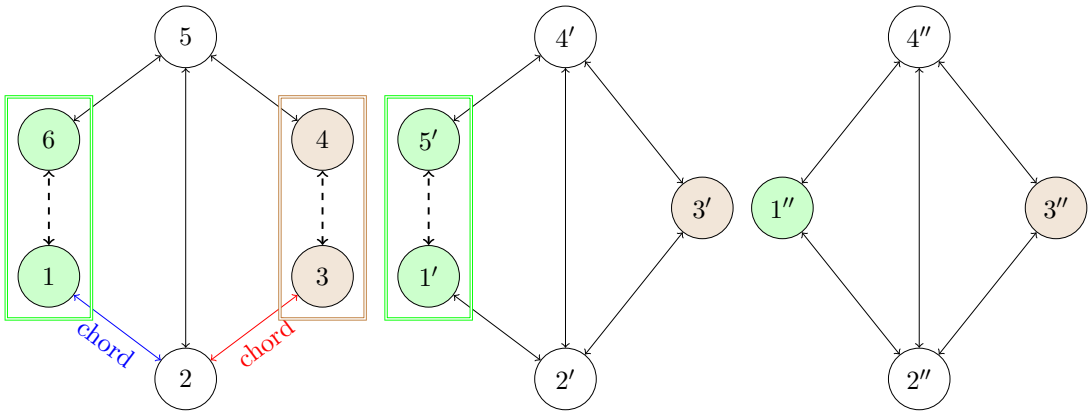


Figure 3.5.: Kinetic diagram for Kinesin from [5]. The 6-state model will be simplified to a 4-state model without changing the number of cycles within the network. In the forward cycle, states 1 and 6 will be merged to state 1", while in the backward cycle, states 3 and 4 are described by state 3" in the coarse-grained network. Except for the fluxes between the merged states, all net-transition fluxes are retained. We eliminated the transition with minimal contribution to entropy production in each fundamental cycle.

will be needed later to calculate fluctuations of steady-state quantities. In equation (B.5) in the supplementary material, it is demonstrated how to calculate the cycle fluxes in the kinesin network. The entropy change after completion of a cycle C can be expressed in terms of the transition affinities within the cycle as defined in equation (2.8). With the choice of chords depicted in figure 2.1, the affinity of the futile cycle C_D can be written as a linear combination of the fundamental cycles as

$$\Delta S_D = \Delta S_F + \Delta S_B. \tag{3.17}$$

The entropy production in a system in a non-equilibrium steady-state can be written in terms of fundamental cycle affinities and the flux through the fundamental cycles defining chords as stated in equation (2.7). This yields

$$\begin{aligned} P &= \sum_l J_l \Delta S_{C_l} \\ &= J_{12}\Delta S_F + J_{23}\Delta S_B. \end{aligned} \tag{3.18}$$

3.2.2. Coarse graining of the kinetic diagram of kinesin without changing the network topology

To simplify the state space, we merge states 3 and 4 in the backward cycle, and subsequently states 1 and 6 in the forward cycle as shown in figure 3.5. The network topology, i.e. the number of fundamental cycles, remains unchanged by these network simplifications. The coarse-grained network still contains 3 cycles, of which 2 are fundamental. By construction of our coarse-graining approach, all transition fluxes are conserved. In particular, the chord fluxes remain unchanged. ($\tilde{J}_{1''2''} = J_{12}$ and $\tilde{J}_{2''3''} = J_{23}$). The flux between the merged nodes is lost however and hence the fundamental cycle affinities are reduced by the transition affinity of the eliminated transition between the merged nodes:

$$\begin{aligned} \Delta\tilde{S}_F &= \Delta S_F - \ln\left(\frac{\alpha_{61}p_6}{\alpha_{16}p_1}\right) \\ \Delta\tilde{S}_B &= \Delta S_B - \ln\left(\frac{\alpha_{34}p_3}{\alpha_{43}p_4}\right) \end{aligned} \tag{3.19}$$

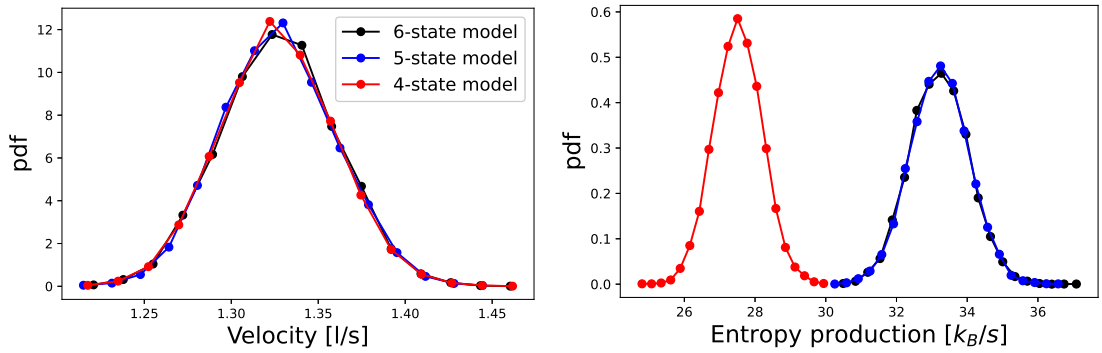
The cycle affinities are retained if the net flux between the merged states is zero. Hence, the coarse graining also reduces the entropy production by

$$\begin{aligned} \Delta P &= P - \tilde{P} \\ &= J_{12}\ln\left(\frac{\alpha_{61}p_6}{\alpha_{16}p_1}\right) + J_{34}\ln\left(\frac{\alpha_{34}p_3}{\alpha_{43}p_4}\right) \end{aligned} \tag{3.20}$$

unless detailed balance is fulfilled between the merged states.

3.2.3. Fluctuations in the kinesin network

Our coarse-graining approach is designed to preserve fluxes, which correspond to averages of observable properties of the kinesin motor. To demonstrate how our coarse-graining approach affects the distribution of quantities such as the entropy production and velocity of kinesin, we simulated 10000 trajectories of finite duration. We compared the original 6-state model of Liepelt and Lipowsky [5] with two coarse-grained systems. In the 5-state system, state 3 and state 4 of the original model are merged such that the backward cycle loses one transition. In the 4-state model, states 1 and 6 are merged in addition, such that the forward cycle loses one transition as well.



(a) (b)

Figure 3.6.: Simulation results for the velocity (a) and the entropy production (b) of a kinesin motor. The rate constants for the 6-state model are from [5] for chemical concentrations $[ATP] = [ADP] = [P] = 1\,\mu M$, stepping size $l = 8\,\text{nm}$ and an external load force $F = 1\,\text{pN}$. 10000 trajectories have been sampled for each model with a simulation time $\tau = 1200\,\text{s}$. The black data corresponds to the original 6-state model from [5] which is depicted in figure 3.5. The blue data represents a coarse-grained system with 5 states, where the original states 3 and 4 have been merged. The four-state model is depicted in red. The entropy production for a trajectory of finite length was calculated as in equation (3.21). The velocity v of a kinesin motor is proportional to the step size l and to the flux along the mechanical transition (J_{25} in the 6-state, $J_{2'4'}$ in the 6-state and $J_{2''4''}$ in the 4-state model). Our coarse-graining approach preserves the mean of the velocity whereas it does not preserve variances of observables. However, the numerical differences for the variance of the velocity are very small.

On the left-hand side in figure 3.6, we show the velocity distribution for the 6-, 5-, and 4-state model obtained from 10000 trajectories with simulation time 1200 s. The variance of the velocity decreases for longer trajectories. The velocity is proportional to the net flux of the mechanical transition $2 \to 5$ in the original model and to the respective mechanical transitions in the coarse-grained models ($2' \to 4'$ in the 5-state model, $2'' \to 4''$ in the 4-state model). Leaving the network topology unchanged, our coarse-graining approach preserves the steady-state flux and hence the mean velocity ($v = \tilde{v} = 1.3361$ l/s) remains unchanged in the course of the coarse-graining procedure. Moreover, the full velocity distribution seems to be unaffected by the coarse graining. But in fact, our coarse-graining approach does not preserve variances of observables. However, by calculating the variance analytically, one can see that the difference and hence the error due to coarse-graining is very small in this case: For the 6-state model, the standard deviation of the velocity is

$\sigma_v = 0.033368526397$ l/s and for the 4-state model $\sigma_{\tilde{v}} = 0.033368526432$ l/s. Coarse graining from 6 to 4 states only leads to a relative increase of the standard deviation of $3 \cdot 10^{-9}$.

The corresponding distributions for the entropy production for trajectories of length 1200 s are shown on the right side of figure 3.6. For a trajectory of duration τ with m transitions $n_0 \to n_1 \to ... \to n_{m-1} \to n_m$ we consider the quantity

$$P_\tau = \frac{k_B}{\tau} \ln \frac{\alpha_{n_1}\alpha_{n_2}...\alpha_{n_{m-1}n_m}}{\alpha_{n_1n_0}\alpha_{n_2n_1}...\alpha_{n_mn_{m-1}}} \tag{3.21}$$

as entropy production as proposed by Lebowitz and Spohn in [18] and also used in e.g. [19]. From equation (3.20), we expect the mean of the entropy production for the coarse-grained systems (4-state model in red and 5-state model in blue) to be reduced compared to the 6-state model (black). The variance is not retained either in the coarse-graining step. However, the numerical differences between the 6-state and the 5-state model are very small, as the edge with the minimal contribution to entropy production in the backward cycle is removed and the backward cycle is dominated by the forward cycle for small loads. Coarse gaining from 6 to 5 states leads to a relative decrease of the mean of the entropy production of about 0.02%. Hence the 6-state and the 5-state model show similar distributions. In general, however, mean and variance of entropy production are not retained by our coarse-graining approach. Coarse graining the edge with minimal contribution in the forward cycle changes the mean entropy production by about 17% compared to the original model. For zero or small external load forces, the forward cycle is dominant and yields the largest contribution to entropy production. With increasing forces, the backward cycle becomes more important. The force dependency will be discussed in more detail in the next section and in figure 4.2, when multi-cyclic and uni-cyclic kinesin models will be compared .

In section B.3 in the supplementary material the variances of the velocity and the entropy production of the kinesin motor are calculated analytically. In figure 4.2, the mean of the velocity and of the entropy production for the 6-state and for the 4-state model is depicted as a function of the external load force. Figure B.2 in the supplementary material shows the variance of the velocity and entropy production distribution as a function of the external load force.

To summarize, the 6-state model for kinesin has been simplified without changing its network topology. The steady-state velocity remains unchanged if the network

topology is preserved. Whereas, the mean of the entropy production always decreases if a transition is eliminated that does not fulfil detailed balance. Variances in general (here for entropy production and velocity) are not preserved.

4. Coarse graining with changed cycle topology

So far, we discussed examples where coarse graining did not change the network topology. In this chapter, we will investigate how characteristic quantities of a system change when the number of fundamental cycles is reduced. First, we consider the molecular motor kinesin in section 4.1 again. Furthermore, we obtain different levels of coarse-grained models of the adaptation network for chemotaxis of Escherichia coli bacteria in section 4.2.

4.1. Iteratively coarse-grained kinesin

4.1.1. General approach

In the previous section, we coarse grained the kinetic diagram for kinesin without changing the cycle topology: Both the 6-state model of Liepelt and Lipowsky [5] and the coarse-grained 4-state model contain three cycles. However, the kinesin literature, in particular earlier studies, also contains models with a single cycle. One prominent example is the uni-cyclic 4-state model by Fisher and Kolomeisky [6]. In this section, we will also allow for the removal of cycles in coarse graining, which allows us to obtain such a uni-cyclic model from the 6-state model with 3 cycles. We use a general approach, in which we iteratively eliminate states from the kinesin model to simplify the description of the molecular motor and to obtain a hierachy of models with different levels of coarse-graining. In each coarse-graining iteration, the two states that are merged are chosen such that the transition or the cycle with minimal entropy production is removed and thus the difference in entropy production between the models is minimal. In general, there will be coarse-graining steps that preserve the cycle topology and steps that change it. For the kinesin model, we can perform four iterations of the coarse-graining step. In each iteration,

the model loses one state, such that we end up with a 2-state network after step 4. The general coarse-graining algorithm is as follows:

1. Find the (chemical) transition (ij) with minimal contribution to the entropy production:

$$\min_{(ij)} J_{ij}\Delta S_{ij} = \min_{(ij)}(\alpha_{ij}p_i - \alpha_{ji}p_j) \cdot \ln\left(\frac{\alpha_{ij}p_i}{\alpha_{ji}p_j}\right) \tag{4.1}$$

2. After finding the transition (ij) with minimal contribution to entropy production, we have to check whether merging states i and j results in a change in cycle topology. If the system loses a cycle C due to merging states i and j, one checks whether the contribution to the entropy production $J_{ij}\Delta S_C$ is smaller than any contribution to the entropy production $J_{mn}\Delta S_{mn}$ for any transition (mn) which is not part of cycle C, i.e. one checks if

$$J_{ij}\Delta S_C < \min_{(mn)\notin C} J_{mn}\Delta S_{mn}. \tag{4.2}$$

3. If equation (4.2) is fulfilled or merging states i and j does not change the cycle topology, merge the pair of states i and j. If equation (4.2) is not fulfilled and merging states m and n does not change the cycle topology, merge states m and n. Otherwise, eliminate the cycle with minimal contribution to the entropy production.

4. Repeat until one obtains an equilibrium two-state network.

We will demonstrate how this algorithm can be used to simplify the model for a kinesin motor under different external load forces. For a given value of the force, the coarse-grained models at different levels of the hierarchy will be compared with respect to their network topology and the distributions of their entropy production and velocity. Kinesin preferably moves in the forward direction when no load force is applied. Its velocity decrease under load [20], the force that results in a zero mean velocity is called stall force (approximately 7 pN [21]). We consider three cases: (a) no load (section 4.1.2), (b) a load below the stall force (section 4.1.3), and (c) a load above the stall force, for which the motor moves backwards (section 4.1.4).

Since the coarse-graining procedure we use generates a hierarchy of models, one can ask which degree of coarse graining is optimal in some sense. It is desirable to

remove as many states as possible, while retaining as much as possible key characteristics of the model. We propose to use the entropy production as the characteristic that should be retained as much as possible. Entropy production is reduced with every step of coarse graining. After four iterations, four states have been eliminated such that kinesin is described by a two-state model, which is by definition in equilibrium and has zero entropy production. To balance the reduction in the number of nodes in the network and the loss of entropy production, we introduce the free energy-like quantity

$$F(\nu) = \frac{\Delta P(\nu)}{P_0} - T \cdot \frac{\Delta N(\nu)}{N_0}. \tag{4.3}$$

Here the change in the number of nodes $\Delta N = N_0 - N(\nu)$ and the change of entropy production $\Delta P = P_0 - P(\nu)$ are both determined with respect to the original 6-state model and normalized to the respective value in the original model denoted with subscript 0. We have lost a fraction $\Delta P(\nu)/P_0$ of the entropy production and eliminated a fraction $\Delta N(\nu)/N_0$ of the states after coarse-graining iteration ν. The weighting factor T controls the relative weight of the information loss (the entropy production) and the simplicity (the number of eliminated states) of our model. In the following, we chose T to be one and hence both terms contribute equally. We can then find the optimal level of coarse-graining by the minimum of $F(\nu)$.

4.1.2. Different levels of coarse-grained models for zero external load force

For a free kinesin motor with zero external force, figure 4.1(a) shows the hierarchy of models at different levels of coarse-graining. A uni-cyclic 3-state model (after 3 iterations) represents a sweet spot between information loss in terms of entropy production and simplicity of the model if both terms are equally weighted ($T = 1$). If no external force is applied, the forward cycle is dominant and yields the largest contribution to the entropy production $P = J_{12}\Delta S_F + J_{23}\Delta S_B$. Therefore, the first two coarse-graining iterations only affect the backward cycle. The first step preserves the network topology and the backward cycle loses the transition with the smallest contribution $J_{ij}\Delta S_{ij}$ to entropy production. This does not affect the two independent transition fluxes J_{12} and J_{23} (chord fluxes). Hence, the mean velocity of the motor is not changed either. The second iteration changes the network topology

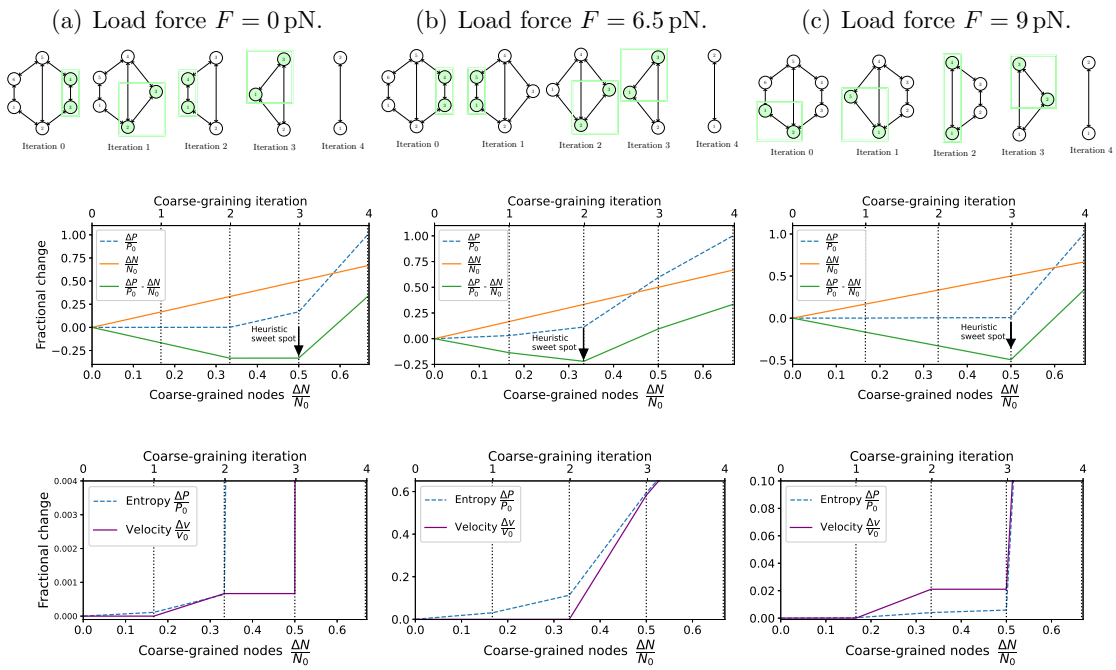


Figure 4.1.: Different models for a kinesin motor. In (a), no external stall force is applied and the forward cycle is dominant. In (b), the kinesin motor is at an external load force of 6.5pN, which is a substantial force, but smaller than the stall force. In (c), the applied load force is larger than the stall force and the backward cycle is dominant.

by eliminating the backward cycle. Thus, the coarse-grained model after steps 2 is a uni-cyclic network. The flux in the uni-cyclic network equals the chord flux J_{12} in the 6-state model. Losing one fundamental cycle and one independent transition flux changes the mean velocity of the motor. Iteration three eliminates a transition within the uni-cyclic 4-state network. The last coarse-graining iteration results in a two-state network which obeys detailed balance by construction. Hence, both the entropy production and the mean velocity of the motor are zero.

In summary, the largest changes in entropy production and mean velocity result from the last coarse-graining step. For zero or small external load forces, the forward cycle is dominant. The backward cycle can be removed by coarse graining, because the approximation error in the steady-state velocity is smaller than 1%.

The uni-cyclic 3-state model represents a sweet spot between information loss and simplicity of the model if the motor works at zero external load force. The differences in the velocity and the entropy production distribution between coarse-grained models and the original model are depicted in figures 4.2 and B.2 for a range of external

load forces to investigate whether the coarse-grained model approximates the original model well for different load forces too. Figure 4.2 shows the mean of the entropy production and the velocity as a function of the external load force, whereas figure B.2 depicts their variance. Both figures indicate that uni-cyclic models approximate the original model well for zero or small external load forces. For increasing forces, the approximation of the velocity and the entropy production distribution becomes less accurate. The next subsection shows that changing the force parameter results in a different hierarchy of models at different levels of coarse graining.

4.1.3. Different levels of coarse-grained models for an external load force smaller than the stall force

With an increasing external load force, the mean velocity of the molecular motor decreases and the forward cycle becomes less dominant in terms of entropy production. Figure 4.1(b) shows different models for a kinesin motor at an external load force of 6.5 pN, which is a substantial force, but smaller than the stall force. The mechanical transition $2 \rightarrow 5$ has the smallest contribution to the steady-state entropy production. We do not merge states 2 and 5 to preserve the biological interpretation of mechanical steps. If we consider only chemical transitions, merging two states in the forward cycle minimises the change in entropy production after the first coarse-graining step. In the second iteration, the backward cycle loses a transition such that the network topology has not been changed after two steps. Therefore, the system still incorporates two fundamental cycles and 2 independent transition fluxes and the mean velocity remains unchanged. The 4-state model after two iterations of coarse-graining is the same model that was already described in section 3.2.2 and shown in figure 3.5. The backward cycle is lost after iteration 3. The change of the network topology results in a change of the mean velocity and a kink in the difference of entropy production with respect to the 6-state model. As before, the last iteration yields a 2-state model which obeys detailed balance by definition and has zero mean velocity and entropy production.

If the change in entropy production and the number of coarse-grained nodes are equally weighted, the sweet spot between information loss and simplicity can be found after coarse-graining iteration 2 as depicted in figure 4.1(b). The resulting tri-cyclic 4-state model preserves the stall force and has an approximation error of 12% for the steady-state entropy production. The uni-cyclic 3-state network after

iteration 3 exhibits pronounced differences in the velocity and entropy production compared to the original 6-state model as depicted in figure 4.2(b). For zero or small external load forces, the changes in these observables are very small for a uni-cyclic model as discussed in the previous section. But with increasing forces acting on the motor, the differences become more pronounced for uni-cyclic models and the tri-cyclic 4-state network becomes the better approximation of the original 6-state model, as it preserves the network topology of the original model. This is an example for how changes in a parameter that modify the transition matrix can affect the results of our coarse-graining algorithm.

4.1.4. Different levels of coarse-grained models for an external load force larger than the stall force

Finally, we consider a force above the stall force. In figure 4.1(c), the 6-state model for the motor is iteratively coarse grained for an external load force $F = 9\,\mathrm{pN}$. Under this condition, the molecular motor moves backwards and the mean velocity is negative. The backward cycle is dominant. The transitions in the forward cycle contribute less to the entropy production than those in the backward cycle. Hence in the first coarse-graining step, a transition in the forward cycle is removed and in the second step, the forward cycle is eliminated entirely. The mean-velocity is changed in those steps, in which the network topology is modified.

4.1.5. Summary

In the previous sections, the 6-state model for kinesin with two fundamental cycles has been iteratively coarse-grained such that the difference in the steady-state entropy production is minimal. In addition, we determined the velocity as a function of force, the main quantity of interest from an experimental point of view. Plotted in figure 4.2(c) and (f), the relative differences between these quantities in coarse-grained models and in the original 6-state models show a pronounced difference between models that preserve the network topology and models that do not. As long as the network topology is preserved, the velocity is unchanged by the coarse-graining and the entropy production is only weakly affected. However, when the network topology is modified in a coarse-graining step, both quantities change. Similar observations are made for the variances of these quantities, as shown in Appendix B.3. The difference between coarse-graining with and without

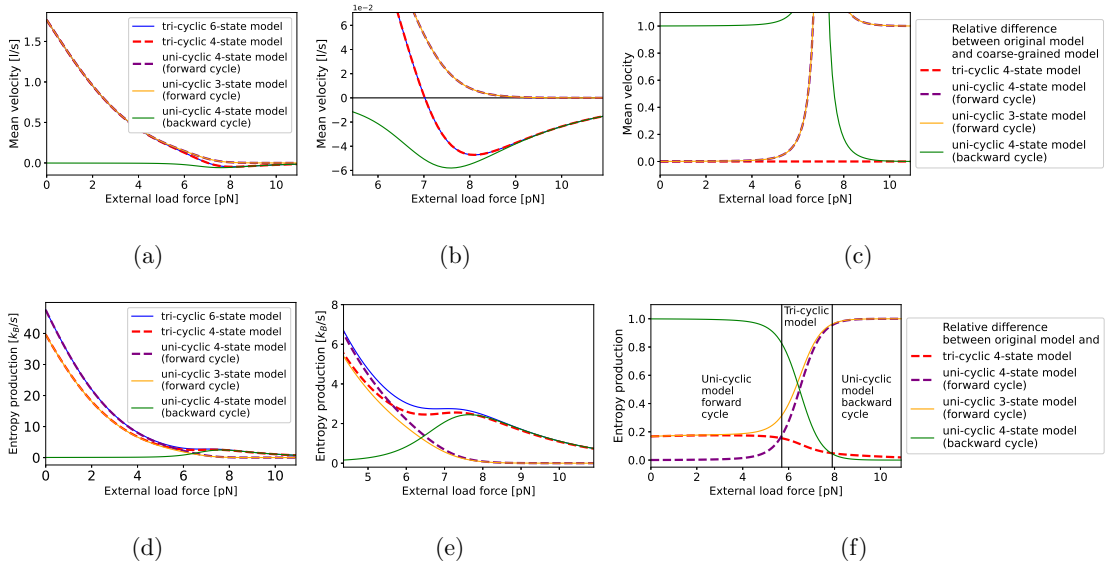


Figure 4.2.: In (a) and (d), the mean velocity and entropy production are plotted against the external load force acting on the motor for different kinesin models, see figure 4.1. The figures in the second column, (b) and (e), show differences between tri-cyclic and uni-cyclic systems for large external load forces. Furthermore, the relative differences between the mean velocity and the entropy production for coarse-grained models and the original 6-state model respectively are depicted in the third column in (c) and (f). The tri-cyclic 4-state model is optimal in terms of preserving the velocity of the motor. Moreover, there are ranges for the load force denoted in (f) to indicate which model is most suitable for approximating the entropy production within the given range. The motor has zero mean velocity for $F = 7\,\mathrm{pN}$. For no or small external load forces, uni-cyclic models resembling the forward cycle approximate the mean velocity and entropy production well with small deviations. While for load forces around the stall force, a coarse-grained model is better, if it has the same network topology as the original model. In figure B.2 in the supplementary material, the variance is plotted against the external load force.

preserved network topology is particularly pronounced for forces around the stall force, where multiple cycles contribute to the dynamics. Uni-cyclic models are a good approximation for load forces that are considerably smaller (or larger) than the stall force. For small forces, the motor walks in the positive direction and has a dominant forward cycle, such that the backward cycle can be neglected. At and above the stall force, however, the backward cycle makes a strong contribution. The uni-cyclic model that resembles the backward cycle approximates the velocity of the motor well if the load force is considerably larger than the stall force and hence the backward cycle is dominant such that the forward cycle can be neglected. We also

note that since changes in network topology affect the velocity, they also shift the stall force itself. The new stall force is $\tilde{F}_{stall} = 13.9\,\mathrm{pN}$. Furthermore, the entropy production becomes zero at the stall force in uni-cyclic networks, while it remains finite in the 3-cycle model, reflecting the fact that in the uni-cyclic model the stall condition is an equilibrium state [17].

Different external load force yield different coarse-graining paths: The forward cycle is dominant for small or no forces. The transitions within the forward cycle have a larger contribution to entropy production than the one in the backward cycle. Therefore, eliminating transitions in the backward cycle results in small changes in entropy production. While for small forces the network topology is changed in the second coarse-graining step, it is only affected in the third step close to the stall force. For forces above the stall force, the backward cycle is dominant and thus, the uni-cyclic model obtained from coarse-graining is based on the backward rather than the forward cycle. In figure 4.2(c) and (f), we can distinguish between load force ranges. For load forces smaller than $5.7\,\mathrm{pN}$, the forward cycle is dominant and the approximation error for the entropy production is minimised by a uni-cyclic model which resembles the forward cycle, cf. figure 4.1(a). For load forces $5.7\,\mathrm{pN} < F < 7.9\,\mathrm{pN}$, the tri-cyclid 4-state model depicted in figures 3.5 and 4.1(b) minimises the difference in entropy production between the original 6-state model and a coarse-grained model. For forces larger $F > 7.9\,\mathrm{pN}$, the backward cycle is dominant and the uni-cyclic model resembling the backward cycle from figure 4.1(c) yields the smallest difference in entropy production with respect to the original model. However, uni-cyclic models do no preserve the stall force in the system and yield to higher approximation errors for the velocity of the motor. Figure 4.2(c) indicates that the tri-cyclic 4-state model is optimal in terms of preserving the velocity of the motor.

Our method can be applied to other systems too. To summarize, we can state that changing the network topology yields larger changes in the distribution of observables than coarse-graining which preserves the network topology if no cycle is dominant. As seen before, the mean and variance of observables can be expressed in terms of fundamental-cycle fluxes. Changing the cycle topology goes along with a change of the number of fundamental cycles which can be seen as a basis of the system in which the behaviour of the system can be expanded. To give an illustrative example, we once again consider the entropy production, which can be calculated

in terms of fluxes and affinities of all transitions (ij) in a system:

$$P = \sum_{(ij)} J_{ij} \Delta S_{ij}$$

If the system contains N states, E transitions, the sum has E terms. Equivalently, the entropy production can be also expressed with a sum with fewer terms, namely $L = E - N + 1$ [equation (2.6)], if we make use of the cycle decomposition. Reducing the number of fundamental cycles L goes along with reducing the number of terms L in the expansion of the entropy production in the fundamental cycles basis. If the network topology is preserved, the chord fluxes J_l are unchanged whereas the cycle affinities ΔS_{C_l} of the fundamental cycles C_l are changed and the number of terms is preserved.

So far, we investigated how coarse graining affects the distribution of the velocity and the entropy production of a kinesin motor. There are further quantities of interest that characterise the mechanical steps of the molecular motor. The walk of kinesin is governed by four dwell time distributions corresponding to the four possible pairs of subsequent forward and backward steps [22] that are called co-steps. The effect of coarse-graining on the dwell time distributions is discussed in section B.4 in the supplementary material. Our coarse-graining approach does not preserve dwell time distributions in general. However, coarse-grained systems with changed network topology exhibit pronounced differences in their dynamics. In section B.4, the dwell time distributions for co-steps and the steady-state probabilities of steps and co-steps are calculated for the original 6-state system and four two coarse-grained systems, namely the 4-state system with three cycles and the uni-cyclic 3-state system from figure 4.1. The first coarse-grained system preserves the number of fundamental cycles. The second one does not. The tri-cyclic coarse-grained model approximates both the steady-state probabilities of steps and co-steps and the dwell time distributions very well for a wide range of load forces.

4.2. Iteratively coarse-grained adaptation network for chemotaxis

In this section, we will discuss how to coarse grain another biochemical system. We will use the adaptation network of chemotaxis of Escherichia coli (E. coli) bacte-

ria to demonstrate how the number of fundamental cycles can be reduced in every coarse-graining step. Again, we will find a sweet spot between simplicity and information loss measured by the cost function in equation (4.3). An Escherichia

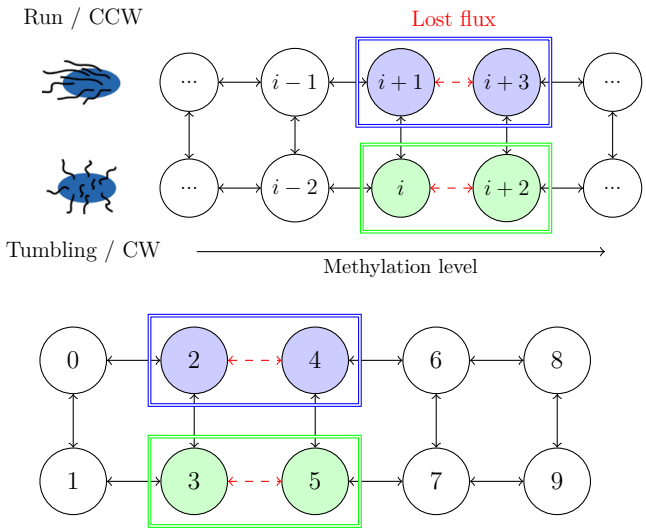


Figure 4.3.: A protein network modulates the tumbling frequency of E. coli [23]. The model for the E. coli chemoreceptor adaptation process is from [24]. Methylation enhances the kinase activity and is responsible for adaptation to changes in the ligand concentration. The states are numbered with i, where even i correspond to active states ($a = 1$) and uneven ones to inactive states ($a = 0$). The activity a affects indirectly the flagella rotation. State i has a methylation level $m = \lfloor i/2 \rfloor$. The time scales for activation and methylation are separated. In each coarse-graining iteration, the blue pair and the green pair of states are merged such that the network loses one fundamental cycle, two nodes, and two edges. The flux along the red-coloured edges is lost due to the coarse-graining procedure. The first figure depicts an arbitrary adaptation network. One can extend the model by increasing the number of methylation states. The second figure shows a model with five different methylation levels and four fundamental cycles. The adaptation model with four fundamental cycles will be coarse grained in 4 steps until two nodes remain in figure 4.4.

coli bacterium has several flagella that all work synchronously. If all flagella rotate counterclockwise, the bacterium moves in an almost straight line. Whereas clockwise rotation causes tumbling and hence, the bacterium changes randomly its direction [25]. The movement is affected by the concentration of nutrients. At low nutrient concentrations, E. coli tumbles more frequently. By switching between clockwise and counterclockwise flagella rotation, E. coli can swim along gradients of nutrients. The switches are generated by a signal transduction network which can be characterized

by three variables [23]. The ligand concentration acts as the input of the system, the average kinase activity is the output and the average methylation level of the receptors acts as the memory of the system. The receptors are distributed all over the membrane and interact with nutrient molecules in the environment. Binding of attractant ligands inhibits the kinase activity. The time scales for ligand binding, kinase response and receptor modification are well separated [26]. The following feedback loop leads to adaptation: A signal s produces a fast response in the output activity a. The change in activity causes a slower change in the negative control element m which may cancel the effect of the stimulus s. We consider a network with two activity states ($a \in \{0, 1\}$) on the vertical axis in figure 4.3 influencing the rotation of the flagella. States with different methylation levels m are depicted on the horizontal axis. The parametrisation of rates is from [24] and is summarized in the supplementary material in section C. If the chords l are chosen to be the edges connecting states with different methylation levels in figure 4.3, the fundamental cycles are squares. The affinity of a fundamental cycle is the ratio of the transition rates in forward and in the backward direction of the cycle. With that, the entropy production of the whole adaptation network can be expressed in terms of the chord fluxes J_l and the affinities of fundamental cycles by

$$P = \sum_l J_l \Delta S_{C_l}. \tag{4.4}$$

We coarse grain the adaptation network by merging two adjacent states with different methylation levels as depicted in figure 4.3. The red-coloured flux between the merged states is lost. The system loses one fundamental cycle. Therefore the entropy production of the coarse-grained system loses one term. Changing the transition rates from merged states to their adjacent states also changes the cycle affinities of the cycles adjacent to the lost cycle. As a concrete example, the difference in entropy production will be explicitly calculated for a network with four fundamental cycles A, B, C and D as depicted in the second diagram in figure 4.3. The difference in entropy production is given by

$$\Delta P = P - \tilde{P} = J_{02}(\Delta S_A - \Delta \tilde{S}_A) + J_{24}\Delta S_B + J_{46}(\Delta S_C - \Delta \tilde{S}_C). \tag{4.5}$$

Coarse graining cycle B changes the affinities of neighbouring cycles A and C. The difference between the cycle affinities in the fine-grained and coarse-grained model

can be expressed as

$$\begin{aligned}\Delta S_A - \Delta \tilde{S}_A &= \ln \frac{\alpha_{32} p_3}{\alpha_{23} p_2} + \ln \frac{\alpha_{23} p_2 + \alpha_{45} p_4}{\alpha_{32} p_3 + \alpha_{54} p_5} \\ \Delta S_C - \Delta \tilde{S}_C &= \ln \frac{\alpha_{45} p_4}{\alpha_{54} p_5} + \ln \frac{\alpha_{32} p_3 + \alpha_{54} p_5}{\alpha_{23} p_2 + \alpha_{45} p_4}.\end{aligned} \tag{4.6}$$

The affinity of cycle D is not affected by coarse-graining cycle B. The difference ΔP in entropy production has three terms, where the term $J_{24}\Delta S_B$ is dominant in the given parametrisation. The dominant term represents the contribution to entropy production of the lost fundamental cycle. The neighbouring cycle affinities are changed by coarse-graining cycle B: The transitions between active and inactive states for the coarse-grained methylation level affect the cycle affinities for the adjacent cycles A and C. Merging adjacent states with different methylation levels does not change the average activity of the bacterium and the average adaptation error if the merged states m and n share the same activity $a_m = a_n \in \{0, 1\}$ as depicted in figure 4.3. The average activity $\langle a \rangle$ is the sum over all probabilities $p_a(m)$ of states with activity $a = 1$ and methylation level m [24] and given by

$$\langle a \rangle = \sum_m p_1(m). \tag{4.7}$$

By construction, all probabilities are retained by our coarse-grain approach. The probabilities of merged states add up such that the activity of the bacterium remains unchanged. The relative adaptation error

$$\epsilon = |1 - \frac{\langle a \rangle}{a_0}|, \tag{4.8}$$

depends on the average activity and hence is unaffected by the coarse grain procedure too. The change in activity a triggers a lower change in the negative control element m, which may bring back activity a to a stimulus-independent level a_0 [24]. We can iterate this coarse-graining procedure. We start with a network with 5 methylation states, hence 10 nodes in total, depicted in figure 4.4. In each coarse-graining step, two edges are cut and the system loses two states and one fundamental cycle. We merge the states such that the change of the entropy production is minimal. After iteration 4, we end up with a two-state model which has no methylation levels anymore and contains only a state with activity $a = 0$ and a state with $a = 1$ (indirectly affecting the running and tumbling behaviour). A two-state model is by

definition in equilibrium and has zero entropy production. In each coarse-graining step, the states will be merged in such a way, that the difference in entropy production is minimal. The largest difference in entropy production P is due to the coarse graining of the last four-node cycle to a two-state system in the last coarse-graining step. The two-state system is in equilibrium and has no more fluxes. Again, it

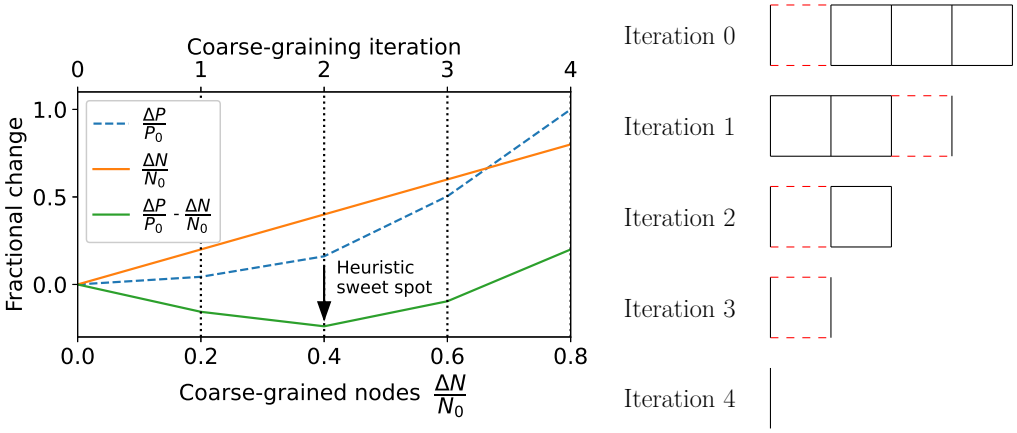


Figure 4.4.: The difference in entropy production ΔP is plotted against the number of coarse-grained nodes ΔN. Both quantities are normed to the respective value of the original network with 10 nodes (and hence 5 methylation levels), which modulates the tumbling frequency of E. coli. To balance the reduction in the number of nodes in the network and the loss of entropy production, we introduce the free energy-like quantity in equation (4.3) as a cost function. We equally weighted the measure for information loss $\Delta P/P_0$ and the measure for simplicity of the model $\Delta N/N_0$. The difference $\Delta P/P_0 - \Delta N/N_0$ is plotted in green and has a minimum indicated by an arrow: After two coarse-graining iterations 40% of all nodes are coarse-grained and the entropy production is reduced by 16%. The dotted edges in the networks depicted on the right side show which edge fluxes are lost in each coarse-graining step and which fundamental cycle yields the minimal contribution to entropy production. The cycle with minimal contribution to entropy production is coarse-grained in each step. The original network has 10 nodes and a basis of four independent cycles. In each coarse-graining iteration, one fundamental cycle is coarse grained. The parameters for the free energy difference $\Delta E(s, m)$ are from [24]. The free energy difference determines the (de-)activation rates in equation (C.1). The dissociation constant is $K_I = 19.2\,\mu\text{M}$ for the inactive receptors and $K_A = 3000\,\mu\text{M}$ for the active receptors. The ligand concentration is $s = 0.1K_I$. The methylation rate $k_{+,0}$ for inactive receptors is $1\,s^{-1}$, while the demethylation rate of active receptors is $k_{-,1} = 0.1\,s^{-1}$.

is desirable to remove as many states as possible, while retaining as much as possible key characteristics of the model. As before, we propose to use the entropy

production as the characteristic that should be retained as much as possible. The entropy production is reduced with every step of coarse-graining. In analogy to the algorithm in section 4.1.1, we can define a coarse-graining algorithm as:

1. Find the cycle C_l with minimal contribution to entropy production:

$$\min_l J_l \Delta S_l,$$

   where the cycle C_l is defined by chord l.

2. Merge two pairs of states that are within the cycle with minimal contribution to entropy as depicted in figure 4.3.

3. Repeat until one obtains an equilibrium two-state network.

To balance the reduction in the number of nodes in the network and the loss of entropy production, we use the free energy-like quantity from equation (4.3)

$$F(\nu) = \frac{\Delta P(\nu)}{P_0} - T \cdot \frac{\Delta N(\nu)}{N_0}.$$

In figure 4.4 with $T = 1$, we can find the minimum of the cost function after iteration 2. We have lost 16% of entropy production and have coarse grained 40% of our nodes. The two-cyclic system in figure 4.4 seems to be optimal in terms of our coarse-graining criterion. It represents a sweet spot between simplicity and information loss. The mean activity and the adaptation error, are unaffected by the coarse-graining procedure.

5. Discussion and Conclusion

We have introduced a coarse-graining procedure for systems governed by a master equation. The approach merges two states such that the probability of the coarse-grained state is the sum of probability of the two fine-grained states by redefining the transition rates according to equation (2.11). All probabilities of states not affected by the merge are retained. The outgoing transition rates of the coarse-grained state are chosen such that the outgoing flux is conserved. This approach changes steady-state probabilities and transition rates only locally. The procedure can be applied iteratively and has no constraints on the network topology. This is in contrast to the coarse-graining approach presented by Altaner and Vollmer[4], which is restricted to merging adjacent bridge states, hence coarse-graining of branches or states with more complicated transitions to the rest of the network - as the one depicted in figure 2.2 - is not possible. The only constraint for our approach is the adjacency of states that shall be merged. As a consequence, the presented approach can (and typically will) reduce the number of fundamental cycles in the systems as shown in section 4.1.

In section 2.5, we minimised the Kullback-Leibler divergence for trajectories to justify our choice of transition rates in the coarse-grained system. The Kullback-Leibler divergence is a special case of the more general α-divergence

$$D_\alpha(q||\tilde{q}) = \frac{1}{\alpha(1-\alpha)} \left\langle 1 - \left(\frac{q}{\tilde{q}}\right)^\alpha \right\rangle_{\tilde{q}} . \tag{5.1}$$

The Kullback-Leibler divergence is not symmetric: $KL(q||\tilde{q}) \neq KL(\tilde{q}||q)$, whereas the α-divergence is symmetric for $\alpha = 1/2$. Minimising the α-divergence for $\alpha = 1/2$ or an arbitrary α could lead to interesting other definitions of transition rates in the coarse-grained system.

Moreover, we propose that for an iterative coarse-graining procedure, balancing the (unwanted) loss of entropy production with the (wanted) reduction of the number nodes of the network can result in a heuristic sweet spot of an optimally coarse-

grained model. The cost function proposed in equation (4.3) quantifies the balance of information loss and simplicity of the model. The relative weight parameter T is so far chosen arbitrarily. With $T = 1$, the information loss (the entropy production) and the simplicity (the number of eliminated states) contribute equally. In the limit of $T \to 0$, no coarse graining at all is favourable and the minimum of the cost function is always the original model. Whereas in the limit of $T \to \infty$, the minimum of the cost function corresponds to a 2-state equilibrium model where the maximal number of states have been merged. While the iterative coarse-graining procedure is not restricted by the features of the network, the criterion for optimal coarse-graining can only be applied to systems out of thermodynamic equilibrium due to the use of the loss in entropy production as a measure of information. Systems in equilibrium have zero entropy production and hence the criterion cannot be used to compare models that are coarse-grained to different levels for equilibrium systems.

Bibliography

[1] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.*, 48, 1976.

[2] G. Hummer and A. Szabo. Optimal dimensionality reduction of multistate kinetic and Markov-state models. *J. Phys. Chem. B*, 119, 2014.

[3] S. Pigolotti and A. Vulpiani. Coarse graining of master equations with fast and slow states. *J. Chem. Phys.*, 128, 2008.

[4] B. Altaner and J. Vollmer. Fluctuation-preserving coarse graining for biochemical systems. *Phys. Rev. Lett.*, 108, 2012.

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

[6] M. E. Fisher and A. B. Kolomeisky. Simple mechanochemistry describes the dynamics of kinesin molecules. *Proc. Natl. Acad. Sci. U.S.A.*, 98(14), 2001.

[7] T.L. Hill. *Free Energy Transduction and Biochemical Cycle Kinetics*. Dover Books on Chemistry. Dover Publications, 2005.

[8] G. Szabo, T. Tomé, and I. Borsos. Probability currents and entropy production in nonequilibrium lattice systems. *Phys. Rev. E*, 82, 2010.

[9] J. M. G. Vilar and J. M. Rubí. Thermodynamics "beyond" local equilibrium. *Proc. Natl. Acad. Sci. U.S.A.*, 98(20), 2001.

[10] 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, 2012.

[11] S. Pressé, K. Ghosh, J. Lee, and K. Dill. Principles of maximum entropy and maximum caliber in statistical physics. *Rev. Mod. Phys.*, 85, 2013.

[12] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Stat.*, 22(1), 1951.

[13] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill. Principles of maximum entropy and maximum caliber in statistical physics. *Rev. Mod. Phys.*, 85, 2013.

[14] A. Kells, V. Koskin, E. Rosta, and A. Annibale. Correlation functions, mean first passage times, and the kemeny constant. *J. Chem. Phys.*, 152(10), 2020.

[15] G. Woehlke and M. Schliwa. Walking on two heads: the many talents of kinesin. *Nat. Rev. Mol. Cell Biol.*, 1(1), 2000.

[16] W. Wang, L. Cao, C. Wang, B Gigant, and M. Knossow. Kinesin, 30 years later: Recent insights from structural studies. *Protein Sci.*, 24(7), 2015.

[17] C. Hyeon, S. Klumpp, and J. N. Onuchic. Kinesin's backsteps under mechanical load. *Phys. Chem. Chem. Phys.*, 11, 2009.

[18] J. L. Lebowitz and H. Spohn. A gallavotti–cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.*, 95(1), 1999.

[19] A. Puglisi, S. Pigolotti, L. Rondoni, and A. Vulpiani. Entropy production and coarse graining in markov processes. *J. Stat. Mech. Theor. Exp.*, 2010(05), 2010.

[20] K. Visscher, M. J. Schnitzer, and S. M. Block. Single kinesin molecules studied with a molecular force clamp. *Nature*, 400(6740), 1999.

[21] N. J. Carter and R. A. Cross. Mechanics of the kinesin step. *Nature*, 435(7040), 2005.

[22] A. Valleriani, S. Liepelt, and R. Lipowsky. Dwell time distributions for kinesin's mechanical steps. *EPL*, 82, 2008.

[23] U. Alon, M. Surette, N. Barkai, and S. Leibler. Robustness in bacterial chemotaxis. *Nature*, 397, 1999.

[24] G. Lan, P. Sartori, S. Neumann, V Sourjik, and Y. Tu. The energy-speed-accuracy trade-off in sensory adaptation. *Nat. Phys.*, 8(5), 2012.

[25] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387, 1997.

[26] Y. Tu, T. S. Shimizu, and H. C. Berg. Modeling the chemotactic response of Escherichia coli to time-varying stimuli. *Proc. Natl. Acad. Sci. U.S.A.*, 105(39), 2008.

[27] P. G. Moschopoulos. The distribution of the sum of independent Gamma random variables. *Ann. Inst. Stat. Math.*, 37(3), 1985.

[28] T. L. Hill and Y. D. Chen. Stochastics of cycle completions (fluxes) in biochemical kinetic diagrams. *Proc. Natl. Acad. Sci. U.S.A.*, 72(4), 1975.

[29] T. Hill. *Free Energy Transduction in Biology: The Steady-State Kinetic and Thermodynamic Formalism.* Academic Press, New York, 1977.

A. Supplementary material for the dwell time distribution in a branch

In this paragraph, the dwell time distribution for the branch of the uni-cyclic network depicted in figure 3.1 will be calculated analytically. The time spent in the branch is the sum over exponentially distributed random variables. If the system starts in state 2, it can either jump to the rest of the system (states 0 or 1) or jump to state 3 and stay in the branch. If the system jumps to state 3, it will be back in state 2 after time t_3. And then again, it can either leave the branch by jumping to states 0 and 1 or remain in the branch by jumping to state 3. Therefore, the waiting time in the branch is the sum of the dwell time t_2 in state 2 and n times the dwell time in state 2 and 3. It is given by

$$t_n = t_2 + n \cdot (t_3 + t_2), n \in \mathbb{N}. \tag{A.1}$$

Starting in state 2, the transition paths are given in general by

$$\begin{aligned} &2 \to (0,1) \\ &2 \to 3 \to 2 \to (0,1) \\ &2 \to 3 \to 2 \to 3 \to 2 \to (0,1) \\ &2 \to (3 \to 2)^n \to (0,1), n \in \mathbb{N}. \end{aligned} \tag{A.2}$$

The branch (states 2 and 3) can be either left, indicated by $\to (0,1)$, or the system can stay in the branch by hopping n times between states 3 and 2, indicated by $(3 \to 2)^n$. Equation (A.1) gives an expression for the time spent in the branch after $n \in \mathbb{N}$ hops between branch states 2 and 3. If the system starts in state 3, the transition paths are almost the same:

$$3 \to 2 \to (3 \to 2)^n \to (0,1). \tag{A.3}$$

The waiting time

$$t_n = t_3 + t_2 + (t_3 + t_2) \cdot n, n \in \mathbb{N},$$

has one additional summand t_3 compared to equation (A.1). To calculate the dwell time distribution for the branch depicted in figure A.1, one has to sum up the distribution of each possible transition path shown in equation (A.2). Each distribution is an n-fold convolution. The distribution of the transition path $2 \to 3$ is a convolution of two exponential distributions:

$$\begin{aligned}
\rho_2(t) &= \lambda_2 e^{-\lambda_2 t} \\
\rho_3(t) &= \lambda_3 e^{-\lambda_3 t} \\
\rho_{2\to3}(t) &= \int_0^t \rho_2(\tau)\rho_3(t-\tau)d\tau \\
&= \frac{\lambda_2\lambda_3}{\lambda_2 - \lambda_3}(e^{-\lambda_3 t} - e^{-\lambda_2 t}).
\end{aligned}$$

The dwell time is a sum over independent, exponentially distributed random variables. The exponential distribution is a special case of the Gamma distribution. A sum of independent variables, which are Gamma distributed, is Gamma-distributed again [27]. To calculate the dwell time distribution explicitly, we will differ between two cases. Firstly, we will consider a system where the sum of all outgoing rates of state 2 and state 3 are equal, i.e. $\lambda_2 = \lambda_3$. Secondly, we will consider arbitrary exit rates λ_2 and λ_3.

Distribution for the dwell time of the branch with same exits rates

The example transition matrix in equation (3.2) yields different exit rates for the branch states 2 and 3. To calculate the dwell time distribution analytically, we consider a special case where the exit rates $\lambda_2 = \lambda_3$ are equal. Therefore, we change α_{32} in equation (3.2) such that

$$A = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 0 & 4 & 0 \end{pmatrix} \text{ and } \tilde{A} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 3 \\ 2/3 & 2/3 & 0 \end{pmatrix}. \tag{A.4}$$

We consider the case, that the starting state is either state 0, 1 or 2. The n-fold convolution for distributions $\rho_2 = \rho_3$ with same rates $\lambda_2 = \lambda_3 = \lambda$ is characterised by the Gamma distribution with density:

$$\begin{aligned}\text{Gamma}(t, \alpha, \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-t\beta} \\ \text{Gamma}(t, 2n+1, \lambda) &= \frac{\lambda^{2n+1}}{(2n)!} t^{2n} e^{-\lambda t},\end{aligned}$$

where the shape parameter α is $2n+1$, because one adds up $2n+1$ random variables, as one can see in equation (A.2). The rate parameter β is the exit rate λ. Let us emphasize that our characterisation of the n-fold convolution requires $\lambda_2 = \lambda_3$ and does not easily generalise to the general setting of $\lambda_2 \neq \lambda_3$. For $n = 0$ one obtains the waiting time distribution $\rho = \lambda e^{-\lambda t}$, which is exponentially distributed. To calculate the waiting distribution $\rho(t)$ for the whole branch, one has to add up all Gamma distributions according to their weight, which is the probability of the transitions

$$\begin{aligned}p(2 \to 3) &= \frac{\alpha_{23}}{\lambda} \\ p(2 \to (0,1)) &= \frac{\alpha_{20} + \alpha_{21}}{\lambda} \\ p(3 \to 2) &= 1.\end{aligned}$$

With this one obtains

$$\begin{aligned}\rho(t) &= \frac{\alpha_{20} + \alpha_{21}}{\lambda} \sum_{n=0}^{\infty} \text{Gamma}(t, 2n+1, \lambda) \left(\frac{\alpha_{23}}{\lambda}\right)^n \\ &= \frac{\alpha_{20} + \alpha_{21}}{\lambda} \sum_{n=0}^{\infty} \frac{\lambda^{2n+1}}{(2n)!} t^{2n} e^{-\lambda t} \left(\frac{\alpha_{23}}{\lambda}\right)^n \\ &= \frac{\alpha_{20} + \alpha_{21}}{\lambda} \lambda e^{-\lambda t} \sum_{n=0}^{\infty} \frac{\lambda^{2n}}{(2n)!} t^{2n} \left(\sqrt{\frac{\alpha_{23}}{\lambda}}\right)^{2n} \\ &= (\alpha_{20} + \alpha_{21}) e^{-\lambda t} \sum_{n=0}^{\infty} \frac{1}{(2n)!} \left(t\sqrt{\alpha_{23}\lambda}\right)^{2n} \\ &= \frac{\alpha_{20} + \alpha_{21}}{2} \left(e^{-(\lambda - \sqrt{\alpha_{23}\lambda})t} + e^{-(\lambda + \sqrt{\alpha_{23}\lambda})t}\right) \\ &= (\alpha_{20} + \alpha_{21}) e^{-\lambda t} \cosh(t\sqrt{\alpha_{23}\lambda}).\end{aligned} \tag{A.5}$$

One can neglect the case that all the transitions start at state 3 as shown in equation (A.3) because state 3 can only be reached via a passage through state 2. Only if

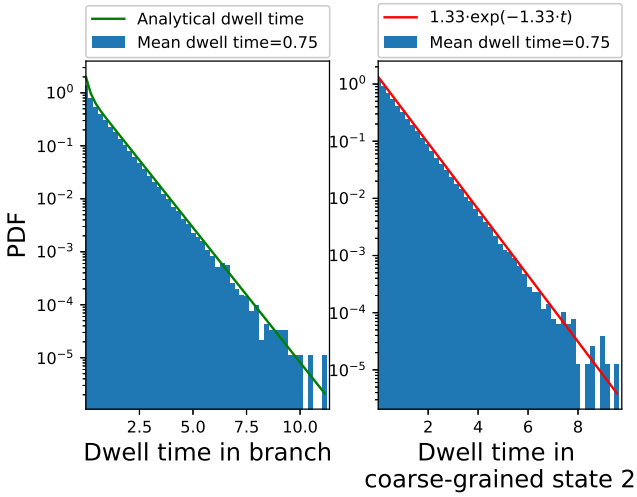


Figure A.1.: Distribution of the time spent in the branch in the original system and in the merged state in the coarse-grained system. The blue histograms are from a Gillespie simulation. Each histogram contains $4 \cdot 10^5$ data points in 100 bins. Both plots describe state 2 and 3 as one state. The first plot refers to the 4-state network with a cycle and a branch depicted in figure 3.1. The simulation differs between branch states 2 and 3. But states 2 and 3 are analyzed as the same state. The second plot shows the exponentially distributed waiting time in state 2 in the coarse-grained picture. The outgoing rates of the new state 2 $\tilde{\alpha}_{20}$ and $\tilde{\alpha}_{21}$ are different than α_{20} and α_{21}. The green line in the first plot refers to equation (A.5). The red line in the right plot is the analytically expected result, the exponential dwell time distribution. The transition rates are the same as in equation (3.2). But $\alpha_{32} = 4$ was changed to make the exit rates of state 2 and 3 equal. This results in the transition rate matrices in equation (A.4). Both distributions have the same mean which is calculated analytically in equation (A.6)

.

the initial condition describes a start at state 3, one has to consider transitions as in equation (A.3) with an even number ($2n$ instead of $2n+1$). The sum over infinitely many Gamma distributions with $\alpha = 2n$ yields equation (A.5) with an sinh instead of cosh. This result is neglectable since it contributes only to the first transition path. After the first transition path, all of them look like the ones depicted in equation (A.2) regardless of the initial condition. The mean of the distribution is given by

$$\begin{aligned} \langle t_2 \rangle &= \int_0^\infty t \cdot \rho(t) = \frac{(\alpha_{20} + \alpha_{21})(\lambda + \alpha_{23})}{\lambda(\lambda - \alpha_{23})^2} \\ &= \frac{\alpha_{23} + \alpha_{32}}{\alpha_{32}(\alpha_{20} + \alpha_{21})}, \end{aligned} \tag{A.6}$$

where we used that $\lambda_2 = \lambda_3 = \lambda$, $\lambda - \alpha_{23} = \alpha_{20} + \alpha_{21}$ and $\lambda = \alpha_{32}$. The distribution of dwell time in the branch has the same mean as the exponential distribution of dwell time for the coarse-grained state 2 which merges the branch states 2 and 3 in the fine-grained system, as given by equation (3.8). The distributions have the same mean even tough the dwell time distribution for the branch is not exponentially distributed.

Arbitrary exit rates: convolution of two gamma distributions

After considering the special case of exit rates $\lambda_2 = \lambda_3$, we will calculate the dwell time distribution of the branch for arbitrary exit rates. We will use that the sum of independent Gamma variables is a Gamma distribution [27]. The waiting time distribution for one transition path is a convolution of two Gamma distributions with different rate parameters , where $n \neq 0$, $\lambda_2 \neq \lambda_3$ and $\lambda_2 < \lambda_3$:

$$\begin{aligned}
\rho_{conv}(n, \lambda_2, \lambda_3, t) =& \text{Gamma}(t, n+1, \lambda_2) * \text{Gamma}(t, n, \lambda_3) \\
=& \int_0^t \text{Gamma}(\tau, n+1, \lambda_2) \cdot \text{Gamma}(t-\tau, n, \lambda_3) d\tau \\
=& \frac{\sqrt{\pi t}}{2 \cdot n!} \lambda_2 (\lambda_2 - \lambda_3)^{1/2-n} e^{-1/2(\lambda_2+\lambda_3)t} (\lambda_2 \lambda_3 t)^n \\
& \cdot \left(I\left(n - \frac{1}{2}, \frac{\lambda_2 - \lambda_3}{2} t\right) - I\left(n + \frac{1}{2}, \frac{\lambda_2 - \lambda_3}{2} t\right) \right),
\end{aligned}$$

where $I(\nu, z)$ describes the modified Bessel function of the first kind with order ν and argument z:

$$\begin{aligned}
I(\alpha, x) &= \sum_{m=0}^{\infty} \frac{1}{m!\,\Gamma(m + \alpha + 1)} \left(\frac{x}{2}\right)^{2m+\alpha} \\
\Gamma(n+1) &= n\Gamma(n) \\
\Gamma\left(n + \frac{1}{2}\right) &= \frac{(2n)!}{n! 4^n} \sqrt{\pi}.
\end{aligned}$$

The total waiting time in the merged state 2 and 3 is given by

$$\rho(t) = \frac{\alpha_{20} + \alpha_{21}}{\lambda_2} \left(\lambda_2 e^{-\lambda_2 t} + \sum_{n=1}^{\infty} \rho_{conv}(n, \lambda_2, \lambda_3, t) \left(\frac{\alpha_{23}}{\lambda_2}\right)^n \right). \tag{A.7}$$

The term $(\alpha_{20} + \alpha_{21})e^{-\lambda_2 t}$ corresponds to the transition path $2 \to (0, 1)$. The term for $n = 1$ corresponds to the transition path $2 \to 3 \to 2 \to (0, 1)$, where the systems hops to branch state 3 before leaving the branch. To calculate the mean of the distribution in equation (A.7), we first compute the mean of $\rho_{conv}(n, \lambda_2, \lambda_3, t)$ first:

$$\int_0^\infty t \cdot \rho_{conv}(n, \lambda_2, \lambda_3, t)dt = \frac{\lambda_3 + (\lambda_2 + \lambda_3) \cdot n}{\lambda_2 \cdot \lambda_3} = \frac{n}{\lambda_3} + \frac{n+1}{\lambda_2}.$$

This allows to write the mean of the waiting time distribution $\rho(t)$ as

$$\langle t_2 \rangle = \int_0^\infty t \cdot \rho(t)dt = \frac{\alpha_{20} + \alpha_{21}}{\lambda_2} \left(\frac{1}{\lambda_2} + \sum_{n=1}^{\infty} \left(\frac{n}{\lambda_3} + \frac{n+1}{\lambda_2} \right) \left(\frac{\alpha_{23}}{\lambda_2} \right)^n \right).$$

One can simplify the sum and calculate the following sums

$$\begin{aligned} \frac{1}{\lambda_3} \sum_{n=1}^{\infty} n \left(\frac{\alpha_{23}}{\lambda_2} \right)^n &= \frac{\alpha_{23}\lambda_2}{\lambda_3(\lambda_2 - \alpha_{23})^2} \\ \sum_{n=1}^{\infty} n \frac{\alpha_{23}^n}{\lambda_2^{n+1}} &= \frac{\alpha_{23}(2\lambda_2 - a)}{(\lambda_2 - a)^2}, \end{aligned} \tag{A.8}$$

which can be evaluated with the geometric series

$$\sum_{n=1}^{\infty} x^n = \frac{x}{1-x}$$

and its derivative

$$\frac{d}{dx} \sum_{n=1}^{\infty} x^n = \sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}.$$

Hence, by equation (A.8) the mean waiting time is given by

$$\langle t_2 \rangle = \frac{\lambda_3 + \alpha_{23}}{\lambda_3(\lambda_2 - \alpha_{23})} = \frac{\alpha_{32} + \alpha_{23}}{\alpha_{32}(\alpha_{20} + \alpha_{21})}, \tag{A.9}$$

where $\lambda_3 = \alpha_{32}$ and $\lambda_2 - \alpha_{23} = \alpha_{20} + \alpha_{21}$. Notably, the distribution of dwell time in the branch has the same mean as the exponential distribution of dwell time for the coarse-grained state 2 which merges the branch states 2 and 3 in the fine-grained system, as given by equation (3.8). The distributions have the same mean even tough the dwell time distribution for the branch is not exponentially distributed.

In general, we do not need the special case with the same exit rates in the branch states ($\lambda_2 = \lambda_3$). The mean of the dwell time distribution of the branch is the same. Considering the special case of $\lambda_2 = \lambda_3$ allowed us to write down the dwell time distribution as a finite sum in equation (A.5). For arbitrary exit rates, the dwell time distribution in equation (A.8) is an infinite sum.

Summary: mean waiting time

We have considered several dwell time distributions for the uni-cyclic systems depicted in figure 3.1. We quickly recap the different cases discussed above: We investigated the mean waiting time

1. in state 2 for the original network with one cycle and one branch,

2. if state 2 and 3 cannot be distinguished, which can be interpreted that state 3 is an internal state of state 2. This is analogous to calculate the dwell time in the branch,

3. for the coarse-grained network (merged state 2 and 3 with new outgoing rates for the merged state).

Table A.1 summarises the dwell time distributions and their respective mean. The waiting time in state 2 is exponentially distributed both in the original and in the coarse-grained network. The waiting time is the reciprocal value of the exit rate λ in the exponential distribution (3.4):

$$\begin{aligned}\langle t_2 \rangle &= \frac{1}{\lambda_2} = \frac{1}{\alpha_{20} + \alpha_{21} + \alpha_{23}} \\ \langle \tilde{t}_2 \rangle &= \frac{1}{\tilde{\lambda}_2} = \frac{\tilde{p}_2}{p_2} \frac{1}{\alpha_{20} + \alpha_{21}} = \frac{\alpha_{23} + \alpha_{32}}{\alpha_{32}(\alpha_{20} + \alpha_{21})}.\end{aligned}$$

The mean waiting time in state 2 is smaller in the original network than in the coarse-grained network as mentioned when discussing equation (3.7). The mean waiting time in the branch states with same exit rates $\lambda_2 = \lambda_3$ is calculated in equation (A.6). The mean waiting time for arbitrary exit rates $\lambda_2 \neq \lambda_3$ is calculated in the previous section in equation (A.9).
To give a descriptive example (see figure A.2) how to calculate the mean waiting time in the branch (state 2 and 3 in figure 3.1), one can calculate the mean first passage time from state 2 or 3 to state 0 or 1 with the method introduced by Hill

	Dwell time distribution	Mean
In state 2	$\rho(t_2) = \lambda_2 \cdot e^{-\lambda_2 t_2}$	$\frac{1}{\lambda_2} = \frac{1}{\alpha_{20} + \alpha_{21} + \alpha_{23}}$
In state 3	$\rho(t_3) = \lambda_3 \cdot e^{-\lambda_3 t_3}$	$\frac{1}{\lambda_3} = \frac{1}{\alpha_{32}}$
Branch ($\lambda_2 \neq \lambda_3$)	equation (A.7)	$\frac{\alpha_{23} + \alpha_{32}}{\alpha_{32}(\alpha_{20} + \alpha_{21})}$
Branch ($\lambda = \lambda_2 = \lambda_3$)	$\rho(t) = (\alpha_{20} + \alpha_{21}) e^{-\lambda t} \cosh(t\sqrt{\alpha_{23}\lambda})$	$\frac{\alpha_{23} + \alpha_{32}}{\alpha_{32}(\alpha_{20} + \alpha_{21})}$
Merged state 2	$\tilde{\rho}(t_2) = \tilde{\lambda}_2 \cdot e^{-\tilde{\lambda}^{cg} t_2}$	$\frac{1}{\tilde{\lambda}_2} = \frac{\alpha_{23} + \alpha_{32}}{\alpha_{32}(\alpha_{20} + \alpha_{21})}$

Table A.1.: The summary of waiting times in states 2, 3 and in the branch is depicted. The original network refers to the system in figure 3.1 with a three-node cycle and a branch. In the coarse-grained system, the states of the branch (2 and 3) are merged. The mean dwell times in the branch are the same although the distributions for the networks are different. The mean waiting time in states of the original and the coarse-grained network is exponentially distributed. The dwell time in the branch is not exponentially distributed if state 2 and 3 are not distinguishable [equations (A.5) and (A.7)]. The mean waiting time in state 2 is smaller in the original network, as given by equation (3.7), than in the ones with the merged state (regardless of introducing new outgoing rates or not). An alternative way to calculate the mean dwell time in the branch is shown in equation (A.11) and figure A.2. The mean first passage time from state 2 to states in the rest of the network (states 0 or 1) yields the same mean dwell time.

in [7]. State 0 and state 1 are considered to be absorbing states. The absorbing flux is given by

$$J_{abs} = (\alpha_{20} + \alpha_{21}) p_2' = \frac{\alpha_{32}(\alpha_{20} + \alpha_{21})}{\alpha_{32} + \alpha_{23}}. \tag{A.10}$$

The primed probability p_2' refers to the probability of being in state 2 in the network, where state 2 is the starting state and state 0 and 1 are absorbing. The mean first passage time is the reciprocal of the absorbing flux and can be expressed as

$$\langle t_2 \rangle = \frac{1}{J_{abs}} = \frac{\alpha_{32} + \alpha_{23}}{\alpha_{32}(\alpha_{20} + \alpha_{21})}, \tag{A.11}$$

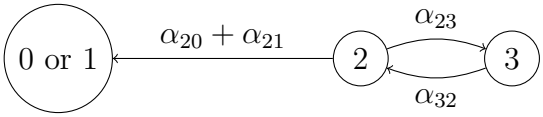


Figure A.2.: The mean waiting time in the state summarizing 2 and 3 can be obtained by calculating the mean first passage time from the branch states 2 and 3 to state 0 or 1 using Hill's method [7]. The mean waiting time is calculated in equation (A.11) and compared to mean dwell time in the branch in table A.1.

which is exactly the result like in the cases discussed before and depicted in table A.1.

Gillespie algorithm

Trajectories can be simulated with a Gillespie algorithm. Relaxation into the steady state or dwell time distributions can be investigated. The exit rate λ_i is the sum over all outgoing rates from state i. First, the exponentially distributed dwell time will be simulated by drawing a random variable z_1 from a uniform distribution between zero and one. The transformation

$$dt = -\frac{1}{\lambda_i} \ln(z_1) \tag{A.12}$$

yields the waiting time in state i before the next transition. To determine which transition will happen, one draws another random variable z_2 and checks in which interval z_2 falls, where

$$I_j = \left[\frac{1}{\lambda_i} \sum_{k=0}^{j-1} a_{ik}, \frac{1}{\lambda_i} \sum_{k=0}^{j} a_{ik} \right). \tag{A.13}$$

If $z_2 \in I_j$ the transition from state i to state j will be simulated. In figures 3.2, 3.3 and A.1, trajectories are simulated with a Gillespie algorithm to compute the dwell time distributions.

B. Supplementary Material for the molecular motor kinesin

B.1. Parametrisation for the molecular motor kinesin

Liepelt and Lipowsky proposed a 6-state model for Kinesin with transition rates

$$\alpha_{ij} = \alpha_{ij}^0 I_{ij}([X])\Phi_{ij}(F), \tag{B.1}$$

where the first factor α_{ij}^0 is a fitted rate constant which is independent of concentrations of $X \in \{\text{ATP, ADP, P}\}$ (see table B.1) or the external load force F [5]. The rate constants are depicted in table B.2 and based on experiments by Visscher et al. [20]. If a transition $i \to j$ does not involve binding of ATP, ADP or P, the second factor $I_{ij}([X])$ is equal to one. Otherwise, it is linear dependent on the concentration of the involved reactant. For instance, ATP is bound during transition $1 \to 2$. Hence, the transition rate α_{12} is linearly dependent on the concentration of ATP: $I_{12}([\text{ATP}]) = [\text{ATP}]$. The force dependence factor $\Phi_{ij}(F)$ is different for the mechanical transitions ($2 \to 5$ and $5 \to 2$) and all other chemical transitions. The

Concentration	$[P]$	$10^{-6}\,\text{M}$
	$[ADP]$	$10^{-6}\,\text{M}$
	$[ATP]$	$10^{-6}\,\text{M}$
Stepping size	l	$8\,\text{nm}$
Thermal energy	k_BT	$4.1 \cdot 10^{-21}\,\text{J}$

Table B.1.: Parameters determining the transition rates for the molecular motor kinesin in equation (B.1) [5].

α^0_{25}	$3 \cdot 10^5\, s^{-1}$
α^0_{52}	$0.24\, s^{-1}$
α^0_{21}	$100\, s^{-1}$
$\alpha^0_{54} = (\alpha^0_{52}/\alpha^0_{25})^2 \cdot \alpha^0_{21}$	
$\alpha^0_{56} = \alpha^0_{61} = \alpha^0_{23} = \alpha^0_{34}$	$200\, s^{-1}$
$\alpha^0_{65} = \alpha^0_{32}$	$0.09\, (\mu\mathrm{Ms})^{-1}$
$\alpha^0_{16} = \alpha^0_{43}$	$0.02\, (\mu\mathrm{Ms})^{-1}$
$\alpha^0_{12} = \alpha^0_{45}$	$1.8\, (\mu\mathrm{Ms})^{-1}$

Table B.2.: Fitted rate constants [20], as given by equation (B.1).

θ	0.3
$\chi_{12} = \chi_{45}$	0.25
$\chi_{23} = \chi_{56}$	0.05
$\chi_{34} = \chi_{61}$	0.05

Table B.3.: Dimensionless load distribution factor θ which governs the mechanical transition rates, as given by equation (B.2). Dimensionless force parameters which govern the force dependence of the chemical transition rates $\chi_{ij} = \chi_{ji}$ [5, 20], cf. equation (B.3).

factors for the mechanical transitions are

$$\begin{aligned} \Phi_{25}(F) &= \exp\left(-\theta \frac{Fl}{k_B T}\right) \\ \Phi_{52}(F) &= \exp\left((1-\theta) \frac{Fl}{k_B T}\right), \end{aligned} \tag{B.2}$$

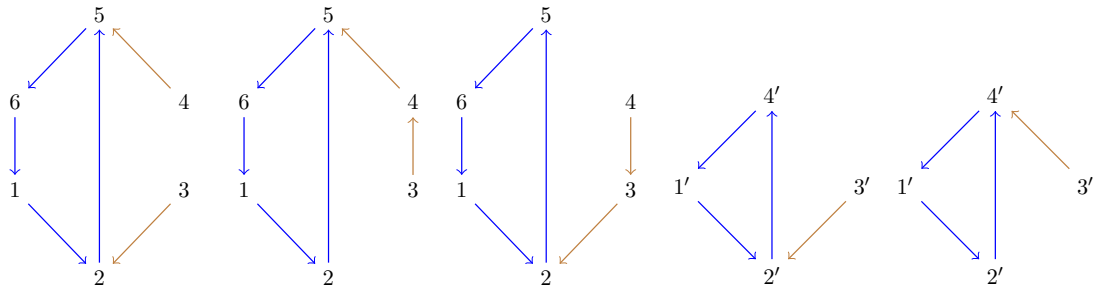
where θ denotes the dimensionless load distribution factor (see table B.3), l the step size, k_B the Boltzmann constant and T the temperature (see table B.1). For all chemical transitions, the force dependence factor

$$\Phi_{ij}(F) = \frac{2}{1 + e^{\chi_{ij} Fl/k_B T}} \tag{B.3}$$

is symmetric because the dimensionless force parameter $\chi_{ij} = \chi_{ji} \geqslant 0$ is symmetric.

B.2. Cycle fluxes in the kinesin network

The cycle fluxes (net numbers of cycle completions per time) can be found from the extension of the diagram method described by Hill in [7]: A flux diagram for a



(a) (b)

Figure B.1.: In (a), the blue arrows contribute to the forward cycle F in the positive (+) direction (counterclockwise). Its algebraic value is $\pi_{F+} = \alpha_{12}\alpha_{25}\alpha_{56}\alpha_{61}$. The brown arrows flow into the forward cycle F. They correspond to the brown-colored terms in equation (B.5). In (b) the diagrams for the forward cycle in plus direction are depicted for a coarse-grained network as shown in figure 3.5. State 1 and 6 in the original network are merged to state 1' in the coarse grained network. States 3 and 4 are merged to state 3'. Again, the brown and blue-colored parts correspond to the color-marked terms in (B.5).

cycle is the cycle itself plus a set of arrows flowing into it as depicted in figure B.1. The algebraic value of a flux diagram for a cycle is the product of rate constants associated with the arrows multiplied by the contribution of the cycle (up to a normalisation constant). The cycle fluxes in the Kinesin network can be calculated with

$$\begin{aligned} J_C &= \frac{\pi_C^+ - \pi_C^-}{S} S_C = J_C^+ - J_C^- \\ J_F &= (\alpha_{45}\alpha_{32} + \alpha_{34}\alpha_{45} + \alpha_{43}\alpha_{32}) \frac{\alpha_{12}\alpha_{25}\alpha_{56}\alpha_{61} - \alpha_{16}\alpha_{65}\alpha_{52}\alpha_{21}}{S} \\ J_B &= (\alpha_{12}\alpha_{65} + \alpha_{61}\alpha_{12} + \alpha_{16}\alpha_{65}) \frac{\alpha_{23}\alpha_{34}\alpha_{45}\alpha_{52} - \alpha_{25}\alpha_{54}\alpha_{43}\alpha_{32}}{S} \\ J_D &= \frac{\alpha_{12}\alpha_{23}\alpha_{34}\alpha_{45}\alpha_{56}\alpha_{61} - \alpha_{16}\alpha_{65}\alpha_{54}\alpha_{43}\alpha_{32}\alpha_{21}}{S}, \end{aligned} \tag{B.4}$$

where the factor S_C is the sum of all possible sets of remaining transitions flowing into the cycle, which flow into the cycle but do not form a cycle themselves. In figure B.1 these transitions are brown colored. The factor S is the sum of the weights of all spanning trees. The cycle fluxes in the coarse-grained 4-state model can be

calculated explicitly as

$$\begin{aligned}
\tilde{J}_C &= \frac{\tilde{\pi}_{C+} - \tilde{\pi}_{C-}}{\tilde{S}} \tilde{S}_C \\
\tilde{J}_F &= (\tilde{\alpha}_{34} + \tilde{\alpha}_{32}) \frac{\tilde{\alpha}_{12}\tilde{\alpha}_{24}\tilde{\alpha}_{41} - \tilde{\alpha}_{14}\tilde{\alpha}_{42}\tilde{\alpha}_{21}}{\tilde{S}} \\
\tilde{J}_B &= (\tilde{\alpha}_{12} + \tilde{\alpha}_{14}) \frac{\tilde{\alpha}_{23}\tilde{\alpha}_{34}\tilde{\alpha}_{42} - \tilde{\alpha}_{24}\tilde{\alpha}_{43}\tilde{\alpha}_{32}}{\tilde{S}} \\
\tilde{J}_D &= \frac{\tilde{\alpha}_{12}\tilde{\alpha}_{23}\tilde{\alpha}_{34}\tilde{\alpha}_{41} - \tilde{\alpha}_{14}\tilde{\alpha}_{43}\tilde{\alpha}_{32}\tilde{\alpha}_{21}}{\tilde{S}}.
\end{aligned} \tag{B.5}$$

B.3. Fluctuations in the kinesin network

Variances of steady-state observables can be explicitly calculated in terms of so-called one-way cycle fluxes [28]. One-way cycle fluxes are not preserved by our coarse-graining method. Hence, variances are retained neither.

The fluctuations of entropy production and velocity can be expressed in terms of fluctuations in the number of completed cycles. All transition fluxes can be written as a linear combination of cycle fluxes as shown in equation (3.16). In equation (B.5), it is demonstrated how to calculate the cycle fluxes in the kinesin network explicitly. The entropy production P and velocity v depend on transition fluxes. In order to calculate the variance of P and v, we need to calculate the variance of transition fluxes, which are linear combinations of cycle fluxes. The latter can be calculated with Hill's diagram method [7]. For a long time interval τ, the net number n_i of completed cycles of type $i \in \{F, B, D\}$ can be treated as an independent random variable which has a Gaussian distribution with mean and variance

$$\begin{aligned}
\bar{n}_i(\tau) &= J_i^{net}\tau \\
\sigma_i^2(\tau) &= (J_{i+} + J_{i-})\tau,
\end{aligned} \tag{B.6}$$

as shown in [29], where $J_{i\pm}$ denotes the cycle flux in positive or negative direction respectively. To avoid confusion, the net-transition flux is denoted as J_i^{net}. But if not stated otherwise all fluxes are net fluxes. The net flux through edge (2,5) can be expressed in terms of cycle fluxes as proposed in equation 3.16, such that

$$J_{25}^{net} = J_F^{net} - J_B^{net}. \tag{B.7}$$

The net numbers of completed forward (F) and backward (B) cycles (depicted in figure 2.1) are independent random variables [29] for long time intervals τ. Thus,

the variance adds up, as the net number of transitions n_{25} is a sum of independent random variables, which have zero covariance:

$$\begin{aligned} n_{25} &= n_F - n_B \\ \bar{n}_{25} &= \bar{n}_F - \bar{n}_B \\ \sigma_{25}^2 &= \sigma_F^2 + \sigma_B^2. \end{aligned} \tag{B.8}$$

The velocity thus has mean and variance

$$\begin{aligned} \frac{\bar{v}}{l} &= J_{25}^{net} = \frac{\bar{n}_{25}}{\tau} \\ \sigma_{v/l}^2 &= \frac{\sigma_{n_{25}}^2}{\tau^2} = \frac{J_{F+} + J_{F-} + J_{B+} + J_{B-}}{\tau}. \end{aligned} \tag{B.9}$$

An explicit expression for the one-way cycle fluxes can be found in equation (B.5). If a coarse-graining iteration preserves the cycle topology (no cycle is lost due to a merged pair of states), the single-cycle fluxes, in general, are not retained. But the sum of net-cycle fluxes (which corresponds to a transition flux) remains unchanged. Thus, the mean of the velocity is retained in all coarse-graining iterations. The variance is not preserved because one-way cycle fluxes are not retained by our coarse-graining approach. Whereas one-way transition fluxes are preserved but cannot be expressed in terms of one-way cycle fluxes. The entropy production

$$\begin{aligned} P &= J_{12}\Delta S_F + J_{23}\Delta S_B \\ &= (J_F + J_D)\Delta S_F + (J_B + J_D)\Delta S_B \\ &= \frac{n_F + n_D}{\tau}\Delta S_F + \frac{n_B + n_D}{\tau}\Delta S_B \end{aligned} \tag{B.10}$$

has variance

$$\begin{aligned} \sigma_P^2 &= (\sigma_{n_F}^2 + \sigma_{n_D}^2)\left(\frac{\Delta S_F}{\tau}\right)^2 + (\sigma_{n_B}^2 + \sigma_{n_D}^2)\left(\frac{\Delta S_B}{\tau}\right)^2 \\ &= \frac{J_{F+} + J_{F-} + J_{D+} + J_{D-}}{\tau}(\Delta S_F)^2 \\ &+ \frac{J_{B+} + J_{B-} + J_{D+} + J_{D-}}{\tau}(\Delta S_B)^2, \end{aligned} \tag{B.11}$$

where we expressed the chord fluxes in terms of cycle fluxes as given by equation (3.16) and used that for long trajectories, the numbers of completed cycles are independent Gaussian distributed random variables as given by equation (B.6) and

[28]. To summarize, the mean of the velocity of the kinesin motor is preserved by the coarse-graining procedure, whereas the mean of the entropy production and variances for entropy production and velocity are not preserved. Figure B.2 shows the variance of the velocity and the entropy production as a function of the external load force acting on the motor for tri-cyclic and uni-cyclic kinesin models.

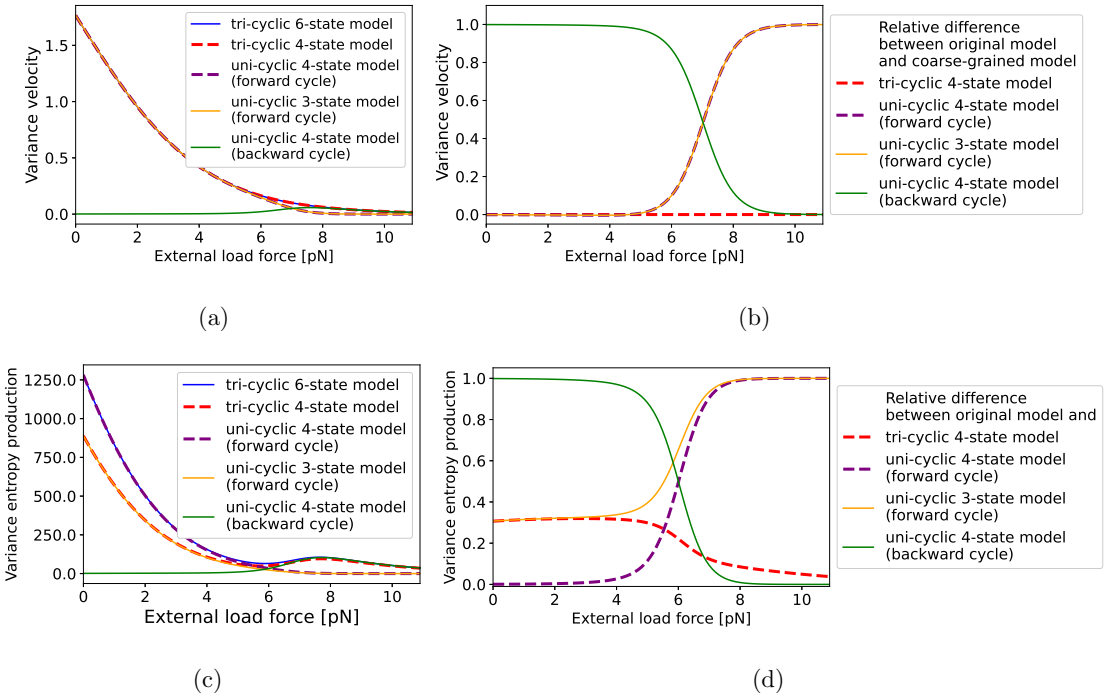


Figure B.2.: The variance of the velocity and the entropy production are plotted against the external load force acting on the motor for different kinesin models. Furthermore, the relative differences between the variance of the velocity and the entropy production for coarse-grained models and the original 6-state model respectively are depicted in the second column. The relative differences in the second column reveal large approximation errors by uni-cyclic systems that resembles the forward cycle for large external load forces. The motor has zero mean velocity for $F = 7\,\mathrm{pN}$. For no or small external load forces, uni-cyclic models approximate the variances for the velocity and the entropy production well with small deviations. While for larger load forces, a coarse-grained model is better, if it has the same network topology as the original model. Both quantities are calculated analytically as described in section B.3. The variance was calculated for trajectories with length $\tau = 1\,\mathrm{s}$.

B.4. Dwell time distributions for kinesin's mechanical steps

So far, we investigated how coarse graining affects the distribution of the velocity and the entropy production of a kinesin motor. We differed between coarse-graining mappings that preserved the network topology of the original model in section 3.2 and between mappings that reduced the number of fundamental cycles in section 4.1. There are further quantities of interest that characterise the mechanical steps of the molecular motor. The walk of kinesin is governed by four dwell time distributions corresponding to the four possible pairs of subsequent forward and backward steps [22] that are called co-steps. Their distributions can be calculated from the master equation if two absorbing states are added. The effective step dynamic is non-Markovian and based on conditional mechanical steps or co-steps [22]. In this section, the dwell time distributions and the steady-state probabilities for the mechanical steps and co-steps in the kinesin model of Liepelt and Lipowsky [5] is discussed.

After introducing the master equation for conditional probabilities and after discussing how to compute dwell time distributions for co-steps in section B.4.1, we compute the steady-state probabilities for mechanical steps and co-steps in a tricyclic 6-state model of the motor kinesin in section B.4.2. To simplify the model, states can be merged as already discussed e.g. in section 3.2.2. In section B.4.3, we compare a coarse-grained 4-state model with the 6-state model with respect to their steady-state probabilities for mechanical steps and co-steps. The 4-state model is further coarse grained in section B.4.4. The 3-state model has a different network topology than the 6-state and the 4-state model. We will discuss the effect of coarse gaining fundamental cycles on the steady-state probabilities for mechanical steps and co-steps. Finally, we will compare the dwell time distributions for the four co-steps and their respective average dwell time for the three different motor models in section B.4.5.

B.4.1. Markov chains with absorbing states

In this paragraph, we will shortly summarize how to compute the distribution of dwell times for Markov chains with absorbing states. In figure B.3(a), states 1-6 are transient states and states 2' and 5' are absorbing states. The conditional

probabilities that the process dwells in state j at time t given that is started in the initial states i at time $t = 0$ are denoted by $P_{ij}(t)$. The time evolution of the probabilities $P_{ij}(t)$ is governed by the master equation [22]

$$\frac{d}{dt}P_{ij} = \sum_{k\neq j}[P_{ik}(t)\alpha_{kj} - P_{ij}(t)\alpha_{jk}]. \tag{B.12}$$

Probability is conserved such that $P_{ii}(0) = 1$ and $\sum_j P_{ij}(t) = 1$. In the steady state (denoted by index st), the probabilities of all transient states are zero and the probabilities of absorbing states add up to one. The steady-state probability P_{ik}^{st} refers to the probability that starting in initial state i the walk is absorbed in absorbing state k. The absorption probabilities P_{ik}^{st} can be computed either by finding the steady state of equation (B.12) or with the so-called closed diagram method presented by Hill [7]. So far, we considered all possible walks of the Markov process. But we are particularly interested in the subset of walks that are eventually absorbed in an absorbing state k. Hence, we consider the conditional probabilities $P_{ij|k}(t)$ that a system dwells at time t in state j given that it started in initial state i and is absorbed in state k. The conditional probability is given by [22]

$$P_{ij|k}(t) = \frac{P_{ij,k}(t)}{P_{ik}^{st}} = \frac{P_{ij}(t)P_{jk}^{st}}{P_{ik}^{st}}. \tag{B.13}$$

The distribution of the dwell times can be obtained by

$$\rho_{i|k}^{ab}(t) = \frac{d}{dt}P_{ik|k}(t) = \frac{1}{P_{ik}^{st}}\frac{d}{dt}P_{ik}(t), \tag{B.14}$$

since $P_{kk}^{st} = 1$ is one for an absorbing state k [22]. To compute the distribution of dwell times analytically, we write the master equation in matrix notation

$$\frac{d\mathbf{P}_i}{dt} = \mathbf{W}\mathbf{P}_i, \tag{B.15}$$

where the subscript i refers to the initial state $i \in \{2, 5\}$ and $(\mathbf{P}_i)_j$ is P_{ij}. The matrix elements of $\mathbf{W}$ are given by

$$W_{mn} = \begin{cases} \alpha_{nm} & \text{for } m \neq n \\ -\sum_{k\neq m}\alpha_{mk} & \text{for } m = n. \end{cases} \tag{B.16}$$

The diagonal elements are called exit rates and ensure conservation of probability, cf. equation (2.21). The exit rate of state m is the sum of all outgoing transition rates. The master equation is solved by

$$\mathbf{P}_i(t) = e^{\mathbf{W}t}\mathbf{P}_i(0), \tag{B.17}$$

where $\mathbf{G}(t) = e^{\mathbf{W}t}$ denotes the propagator.

B.4.2. Steady-state probabilities for mechanical steps and co-steps in a tri-cyclic kinesin model

We will consider four different co-steps that correspond to forward-after-forward steps (ff), backward-after-forward steps (fb), forward-after-backward steps (bf) and backward-after-backward steps (bb) [22]. In a network with absorbing states, the mean time before (any) absorption can be calculated via Hill's approach [7]. Figure B.3 shows the 6-state model for kinesin [5]. The red-coloured mechanical transitions lead to absorbing states. The transition $2 \to 5'$ corresponds to a mechanical forward step, the transition $5 \to 2'$ corresponds to a mechanical backward step. All other transitions are chemical and involve binding or unbinding of ADP or ATP. If the motor walked forward in the previous step, state 5 is occupied. If it walked backwards, state 2 is occupied. The mean time before absorption and the absorbing flux can be calculated if the transitions to absorbing states are redirected to the starting state [7]. The resulting diagram is called a closed diagram. The steady-state distribution of the closed diagram yields the absorption flux. In figure B.3(b), the system starts in state 5 (after a previous forward transition). Hence, the absorbing fluxes are redirected to the initial state 5. The absorbing flux is given by

$$J_{abs} = J_{ff} + J_{fb} = \alpha_{25'}p'_2 + \alpha_{52'}p'_5, \tag{B.18}$$

where the primed probability refers to the closed diagram in figure B.3(b) with initial state 5. Given that the system starts in state 5, the probability, that the motor moves forward ($2 \to 5'$) and hence performs a forward-after-forward step (ff), can be expressed as

$$M_{ff} = P^{st}_{55'} = \frac{J_{ff}}{J_{ff} + J_{fb}} = \frac{\alpha_{25'}p'_2}{\alpha_{25'}p'_2 + \alpha_{52'}p'_5}, \tag{B.19}$$

where $P^{st}_{55'}$ refers to the steady-state of the mater equation (B.12). Alternatively, the motor can walk backwards ($5 \to 2'$) after a forward step. The probability for a forward-backward step (fb) is given by

$$M_{fb} = P^{st}_{52'} = \frac{J_{fb}}{J_{ff} + J_{fb}} = \frac{\alpha_{52'} p'_5}{\alpha_{25'} p'_2 + \alpha_{52'} p'_5} = 1 - M_{ff}. \tag{B.20}$$

If the motor walked backwards in the previous step, the system is in state 2. Figure B.3(c) shows the closed diagram for a system starting in state 2 (after a backward step). The absorbing flux is governed by

$$J_{abs} = J_{bf} + J_{bb} = \alpha_{25'} p'_2 + \alpha_{52'} p'_5, \tag{B.21}$$

where the primed probability refers to the closed diagram in figure B.3(c) with initial state 2. The probability that state 5' is absorbing given that the system started in state 2 is given by

$$M_{bf} = P^{st}_{25'} = \frac{J_{bf}}{J_{bf} + J_{bb}} = \frac{\alpha_{25'} p'_2}{\alpha_{25'} p'_2 + \alpha_{52'} p'_5}, \tag{B.22}$$

and is equal to the probability of a forward step after a backward step. The primed probabilities refer to the steady-state probabilities in the closed diagram in figure B.3(c). The probability for a backward step after a backward step can be expressed as

$$M_{bb} = P^{st}_{22'} = \frac{J_{bb}}{J_{bf} + J_{bb}} = \frac{\alpha_{52'} p'_5}{\alpha_{25'} p'_2 + \alpha_{52'} p'_5} = 1 - M_{bf}. \tag{B.23}$$

The absorption probabilities M_ν ($\nu \in \{ff, fb, bf, ff\}$) define a new discrete-time Markov chain that represents the sequence of forward (f) and backward (b) steps of the motor [22]. The recursion

$$(\pi'_f, \pi'_b) = (\pi_f, \pi_b) \begin{pmatrix} M_{ff} & M_{fb} \\ M_{bf} & M_{bb} \end{pmatrix} \tag{B.24}$$

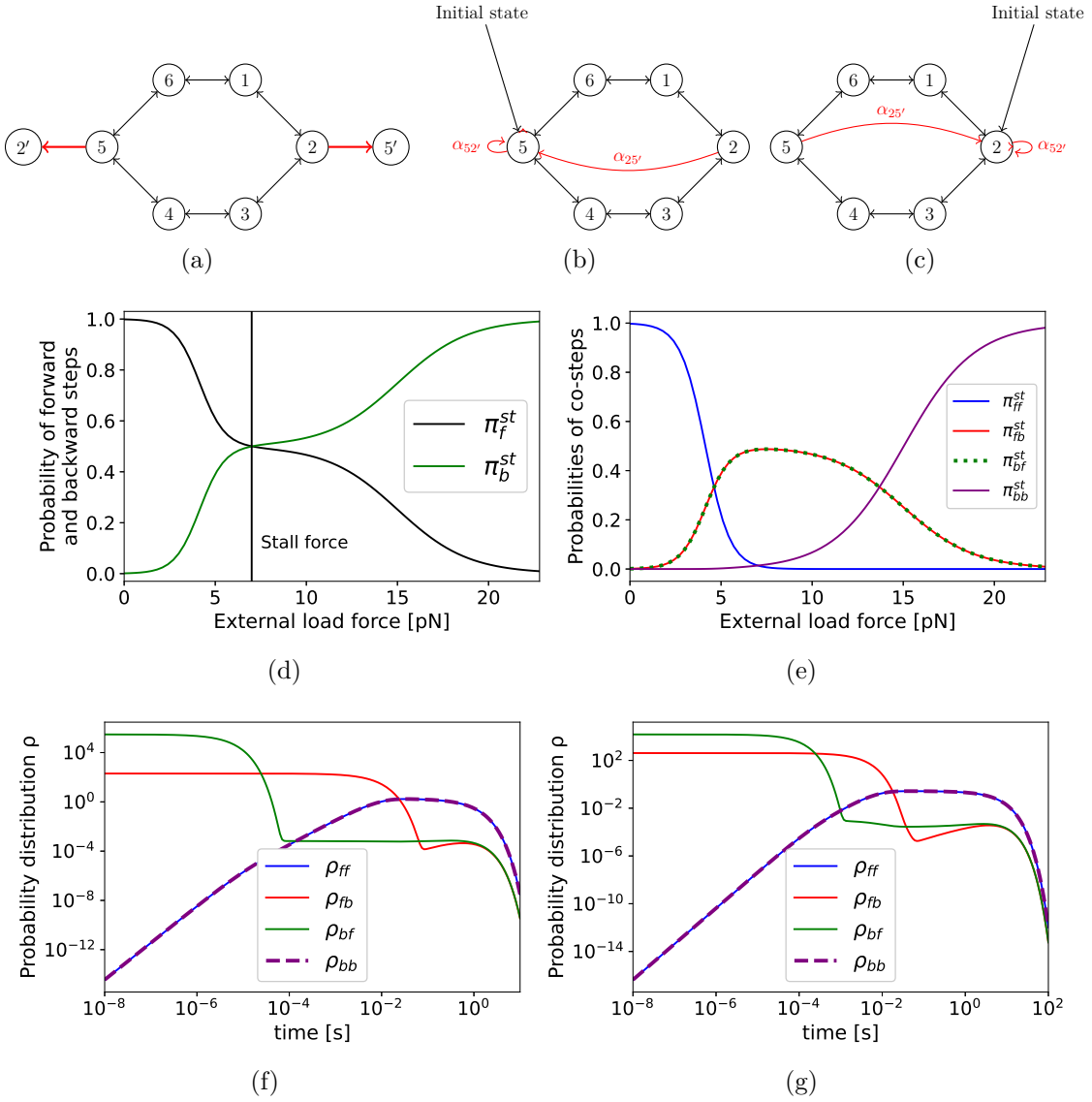


Figure B.3.: 6-state model for the kinesin motor with absorbing states. In (a), states 2' and 5' are absorbing states. The closed diagram in (b) allows to calculated the absorbing flux given that the system starts in state 5 [7]. Starting in state 5 corrseponds to a previous mechanical forward transtion. If the motor walked backwards before, state 2 is the starting state and the mechanical transitions would be redirected to state 2. The closed diagram for a system starting in state 2 (after a backward step) is depicted in (c). In (d), the steady-state probabilities for forward and backward steps from equation (B.25) are plotted as a function of the external load force. In (e), the probabilities of co-steps from equation (B.26) are plotted against the external load force. The co-steps distributions are depicted for zero load force in (f) and for $F = 5\,\mathrm{pN}$ in (g). The paramtrisation is from [5] with concentrations $[ATP] = [ADP] = [P] = 1\,\mu\mathrm{M}$.

for the probability π_f to make a forward step and the probability π_b to make a backward step has a steady state given by [22]

$$\begin{aligned} \pi_f^{st} &= \frac{M_{bf}}{M_{bf} + M_{fb}} \\ \pi_b^{st} &= \frac{M_{fb}}{M_{bf} + M_{fb}} = 1 - \pi_f^{st}. \end{aligned} \tag{B.25}$$

The steady-state probabilities for a forward or backward step are depicted as a function of the external load force in figure B.3(d). The probabilities for a forward and a backward step are equal for the stall force $F_{Stall} = 7\,\mathrm{pN}$. For small external load forces, the forward step is dominant. For large external load force, the backward step dominates. The steady-state probabilities of the four co-steps (ff, fb, bf, bb) are given by [22]

$$\begin{aligned} \pi_{ff}^{st} &= \pi_f^{st} M_{ff} = \frac{M_{bf} M_{ff}}{M_{bf} + M_{fb}} \\ \pi_{fb}^{st} &= \pi_f^{st} M_{fb} = \frac{M_{bf} M_{fb}}{M_{bf} + M_{fb}} \\ \pi_{bf}^{st} &= \pi_b^{st} M_{bf} = \frac{M_{fb} M_{bf}}{M_{bf} + M_{fb}} \\ \pi_{bb}^{st} &= \pi_b^{st} M_{bb} = \frac{M_{fb} M_{bb}}{M_{bf} + M_{fb}}. \end{aligned} \tag{B.26}$$

Figure B.3(e) shows the probabilities of the co-steps (ff, fb, bf, bb) as a function of the external load force. The co-steps forward-after-forward (ff) and backward-after-backward (bb) dominate for forces smaller and larger than the stall force $F_{Stall} = 7\,\mathrm{pN}$. Close to the stall force, the co-steps bf and fb are most probable. A backward-after-forward and a forward-after-backward step does not lead to a net displacement. For small forces, the motor makes primarily forward-after-forward steps. For forces close to the stall force, the motor alternates between forward and backward steps. And for large forces, the motor primarily makes backward-after-backwards steps. The dwell time distributions of the co-steps can computed according to equation

(B.14) such that

$$\begin{aligned}
\rho_{ff}(t) &= \frac{\dot{P}_{55'}(t)}{P^{st}_{55'}} = \frac{P_{52}(t)\alpha_{25'}}{M_{ff}} \\
\rho_{bf}(t) &= \frac{\dot{P}_{25'}(t)}{P^{st}_{25'}} = \frac{P_{22}(t)\alpha_{25'}}{M_{bf}} \\
\rho_{fb}(t) &= \frac{\dot{P}_{52'}(t)}{P^{st}_{52'}} = \frac{P_{55}(t)\alpha_{52'}}{M_{fb}} \\
\rho_{bb}(t) &= \frac{\dot{P}_{22'}(t)}{P^{st}_{22'}} = \frac{P_{25}(t)\alpha_{52'}}{M_{bb}},
\end{aligned} \tag{B.27}$$

where the time derivative $\dot{P}_{ik'}(t)$ of the probability that the motor is in state k' under the condition it started in initial state i is given by the master equation (B.12). For an absorbing state k', the outflux is zero such that $\dot{P}_{ik'}(t) = P_{ij}w_{jk'}$. The dwell time distributions can be computed by integrating the master equation. The dwell time distributions for the co-steps in the original system is depicted for two values for the external load force in figure B.3(f) and (g). All four dwell time distributions for the co-steps are nonexponential. The non-exponential distributons demonstrate the non-Markovian character of the mechanical stepping process [22]. The dwell time distributions $\rho_{ff}(t)$ and $\rho_{bb}(t)$ for forward-after-forward and backward-after-backward steps decrease to zero for small dwell times and are identical. To make the same type of step again, the motor has to visit at least two other states (either states 1 and 6 or states 3 and 4) as shown in figure B.3(a). The transition rates proposed by Liepelt and Lipowsky [5] exhibit a symmetry: $\alpha_{23} = \alpha_{56}$, $\alpha_{34} = \alpha_{61}$, $\alpha_{45} = \alpha_{12}$, $\alpha_{32} = \alpha_{65}$ and $\alpha_{43} = \alpha_{16}$, cf. section B.1. The symmetry implies that the dwell times for the motor in the states 6 and 1 are equal to those on states 3 and 4, respectively. As a result, the dwell time distributions for ff and bb are identical $\rho_{ff}(t) = \rho_{bb}(t)$ [22]. The dwell distributions ρ_{fb} and ρ_{bf} for forward-after-backward and for backward-after-forward steps exhibit several characteristic time scales. Both have a boundary maximum at vanishing dwell times corresponding to two subsequent mechanical transitions without any intervening chemical transition [22]. They exhibit a second shoulder at larger dwell times that correspond to cycles of intervening states chemical transitions that start from state 2 or state 5 and return to the same state before making the next mechanical transition [22].

B.4.3. Steady-state probabilities for mechanical steps and co-steps in a coarse-grained tri-cyclic kinesin model

The 6-state model for the kinesin motor can be coarse grained without changing its network topology as already discussed in section 3.2. The coloured states in the 6-state model in figure B.4(a) are merged. For the following calculations, the coarse-graining order is not important. It does not matter which pair of states is merged first in figure B.4(a) because the coarse-grained transition rates are unaffected by the coarse-graining order as long as the mapping between states in the original and coarse-grained system is the same. We calculate the probabilities of the co-steps (ff, fb, bf, bb) for the coarse-grained 4-state system depicted in figure B.4(e) as before. The closed diagram in figure B.4(c) allows us to calculate the absorption probabilities given that the system walked in the forward direction in a previous step and hence starts in state 4. The absorbing transitions are redirected to the starting state. The absorbing flux is given by

$$\tilde{J}_{abs} = \tilde{J}_{ff} + \tilde{J}_{fb} = \tilde{\alpha}_{24'}\tilde{p}'_2 + \tilde{\alpha}_{42'}\tilde{p}'_4, \tag{B.28}$$

where the primed probabilities refer to the steady-state probabilities in the closed system in figure B.4(c). Quantities with a tilde refer to the coarse-grained system. Given that the system starts in state 4, the probability, that the motor moves forward and hence performs a forward-after-forward step (ff), can be expressed as

$$\tilde{M}_{ff} = \frac{\tilde{J}_{ff}}{\tilde{J}_{ff} + \tilde{J}_{fb}} = \frac{\tilde{\alpha}_{24'}\tilde{p}'_2}{\tilde{\alpha}_{24'}\tilde{p}'_2 + \tilde{\alpha}_{42'}\tilde{p}'_4}. \tag{B.29}$$

The probability for a forward-backward step (fb) is given by

$$\tilde{M}_{fb} = \frac{\tilde{J}_{fb}}{\tilde{J}_{ff} + \tilde{J}_{fb}} = \frac{\tilde{\alpha}_{42'}\tilde{p}'_4}{\tilde{\alpha}_{24'}\tilde{p}'_1 + \tilde{\alpha}_{42'}\tilde{p}'_4} = 1 - \tilde{M}_{ff}. \tag{B.30}$$

If the motor walked backwards, the system is in state 2. We can define the closed diagram for a system starting in state 2 (after a backward step) and calculate the absorption probabilities $\tilde{M}_{bf}$ and $\tilde{M}_{bb}$ analogously. The steady state probabilities $\tilde{\pi}_f$

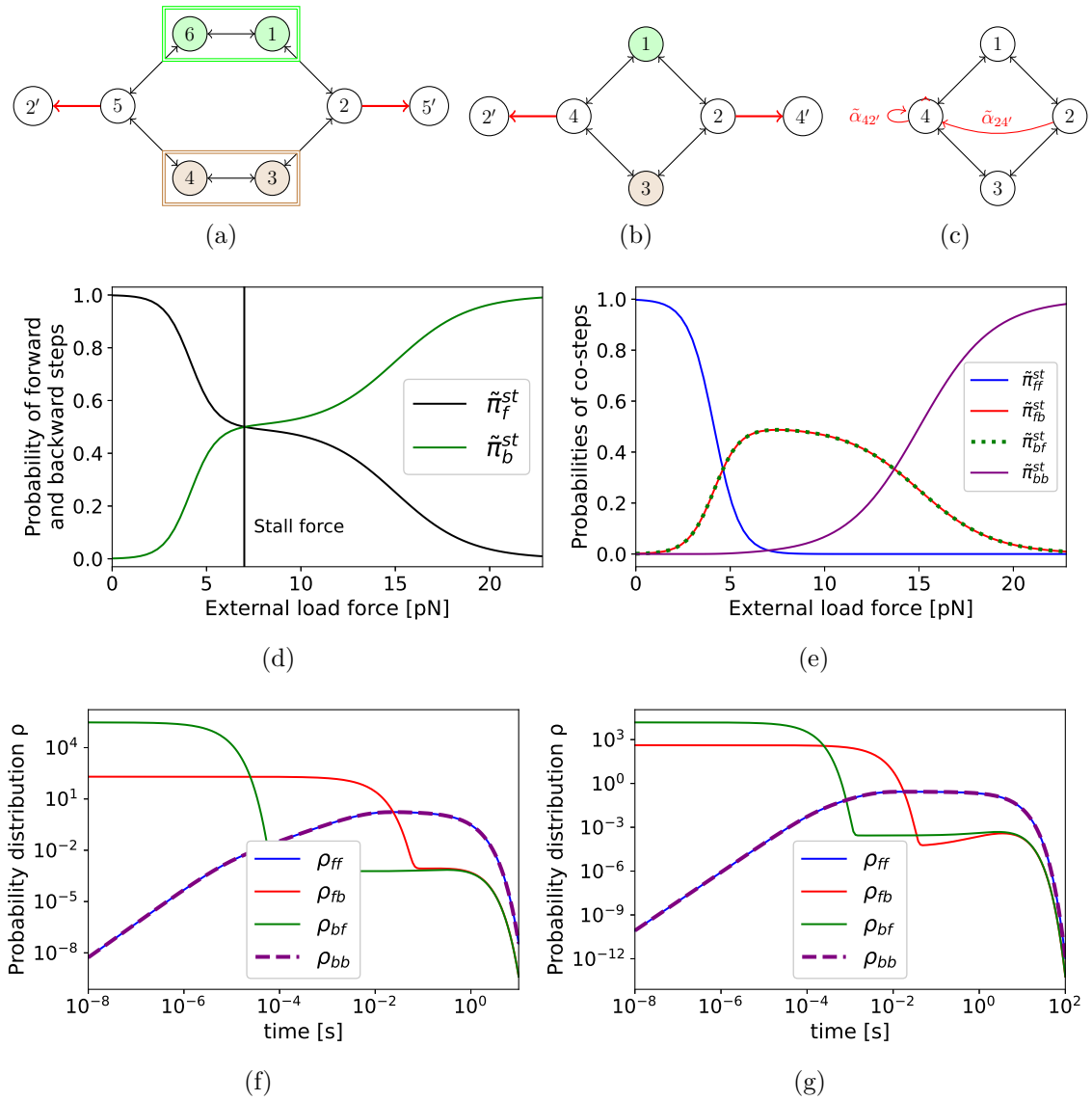


Figure B.4.: The green-coloured and brown-coloured states in the 6-state model in (a) are merged. The coarse-grained system for the kinesin motor is depicted in (b). Again, the absorbing mechanical transition rates are coloured in red. The closed diagram (c) allows the calculation of the absorption flux according to Hill's method [7] given that the system performed a forward step before and hence starts in state 4. The absorbing fluxes (coloured in red) are redirected to the initial state. The closded diagram for initial state 2 is not depicted. In (d), the probabilities for forward and backward steps from equation (B.31) are depicted. In (e), the probabilities of co-steps in the coarse-grained system from equation (B.32) are plotted against the external load force. The characteristic stepping behaviour of the 6-state model depicted in figure B.3 is retained and the tri-cyclic 4-state model approximates both the steady-state probabilities of steps and co-steps very well for a wide range of load forces. The co-steps distributions are depicted for zero load force in (f) and for $F = 5\,\mathrm{pN}$ in (g).

to make a forward step and $\tilde{\pi}_b$ to make a backward step are governed by

$$\begin{aligned}\tilde{\pi}_f^{st} &= \frac{\tilde{M}_{bf}}{\tilde{M}_{bf}+\tilde{M}_{fb}} \\ \tilde{\pi}_b^{st} &= \frac{\tilde{M}_{fb}}{\tilde{M}_{bf}+\tilde{M}_{fb}} = 1-\tilde{\pi}_f^{st},\end{aligned} \tag{B.31}$$

in analogy to equation (B.25). The probabilities for a forward or a backward step are plotted as a function of the external load force in figure B.4(d). The probabilities $\tilde{\pi}_f$ and $\tilde{\pi}_b$ are equal if the stall force F_{Stall} prevents the motor from any net displacement. Merging states 1 and 6 and states 3 and 4 does not change the cycle topology of the network and hence does not alter the stall force. As before, the forward step dominates for small forces, whereas the backward step dominates for large forces. In the coarse-grained system, the steady-state probabilities of the four co-steps (ff, fb, bf, bb) are given by [22]

$$\begin{aligned}\tilde{\pi}_{ff}^{st} &= \tilde{\pi}_f^{st}\tilde{M}_{ff} = \frac{\tilde{M}_{bf}\tilde{M}_{ff}}{\tilde{M}_{bf}+\tilde{M}_{fb}} \\ \tilde{\pi}_{fb}^{st} &= \tilde{\pi}_f^{st}\tilde{M}_{fb} = \frac{\tilde{M}_{bf}\tilde{M}_{fb}}{\tilde{M}_{bf}+\tilde{M}_{fb}} \\ \tilde{\pi}_{bf}^{st} &= \tilde{\pi}_b^{st}\tilde{M}_{bf} = \frac{\tilde{M}_{fb}\tilde{M}_{bf}}{\tilde{M}_{bf}+\tilde{M}_{fb}} \\ \tilde{\pi}_{bb}^{st} &= \tilde{\pi}_b^{st}\tilde{M}_{bb} = \frac{\tilde{M}_{fb}\tilde{M}_{bb}}{\tilde{M}_{bf}+\tilde{M}_{fb}}.\end{aligned} \tag{B.32}$$

Figure B.4(e) shows the probabilities of the co-steps (ff, fb, bf, bb) as a function of the external load force. Comparing figures B.3(e) and B.4(e) indicates that the steady-state probabilities are well approximated in the coarse-grained system. The maximal approximation error is of order 10^{-6}. We can conclude that the tri-cyclic 4-state model approximates both the steady-state probabilities of steps and co-steps very well for a wide range of load forces. The co-steps in the tri-cyclic coarse-grained

4-state model are distributed according to

$$\begin{aligned}
\tilde{\rho}_{ff}(t) &= \frac{\dot{\tilde{P}}_{44'}(t)}{\tilde{P}^{st}_{44'}} = \frac{\tilde{P}_{42}(t)\tilde{\alpha}_{24'}}{\tilde{M}_{ff}} \\
\tilde{\rho}_{bf}(t) &= \frac{\dot{\tilde{P}}_{24'}(t)}{\tilde{P}^{st}_{24'}} = \frac{\tilde{P}_{22}(t)\tilde{\alpha}_{24'}}{\tilde{M}_{bf}} \\
\tilde{\rho}_{fb}(t) &= \frac{\dot{\tilde{P}}_{42'}(t)}{\tilde{P}^{st}_{42'}} = \frac{\tilde{P}_{44}(t)\tilde{\alpha}_{42'}}{\tilde{M}_{fb}} \\
\tilde{\rho}_{bb}(t) &= \frac{\dot{\tilde{P}}_{22'}(t)}{\tilde{P}^{st}_{22'}} = \frac{\tilde{P}_{24}(t)\tilde{\alpha}_{42'}}{\tilde{M}_{bb}}.
\end{aligned} \tag{B.33}$$

Figures B.4(f) and (g) show the dwell time distributions for the co-steps in the coarse-grained tri-cyclic system for two values for the external load force. Remarkably, the distributions for two steps of the same kind (bb and ff) are identical as in the 6-state model. Merging states 3 and 4 in the backward cycle and states 1 and 6 in the forward cycle preserves the symmetry in the transition rates in the original model. The pair of transition rates $\tilde{\alpha}_{23} = \tilde{\alpha}_{41}$ in the coarse-grained model corresponds to $\alpha_{23} = \alpha_{56}$. The transition $2 \to 3$ in the backward cycle still has the same rate as the respective transition $4 \to 1$ in the backward cycle. There are two other pairs of transitions with identical rates: $\tilde{\alpha}_{32} = \tilde{\alpha}_{14}$ and $\tilde{\alpha}_{34} = \tilde{\alpha}_{12}$. Furthermore, the distributions for alternating steps ρ_{fb} and ρ_{bf} exhibit two shoulders as in the original model, cf. figures B.3(f) and (g). The distributions for the co-steps in the original and the coarse-grained model are compared in section B.4.5.

B.4.4. Steady-state probabilities for mechanical steps and co-steps in a coarse-grained uni-cyclic kinesin model

The 6-state model for the kinesin motor can be coarse grained such that only one cycle remains as discussed in section 4.1.1. The coloured states in the 6-state model in figure B.5(a) are merged. We calculate the probabilities of the co-steps (ff, fb, bf, bb) for the coarse-grained uni-cyclic 3-state system depicted in B.5 as before. The closed diagram in figure B.5(c) allows us to calculate the absorption probabilities given that the system walked in the forward direction in a previous step and hence starts in state 3. The absorbing flux is given by

$$\tilde{J}_{abs} = \tilde{J}_{ff} + \tilde{J}_{fb} = \tilde{\alpha}_{23'}\tilde{p}'_2 + \tilde{\alpha}_{32'}\tilde{p}'_3, \tag{B.34}$$

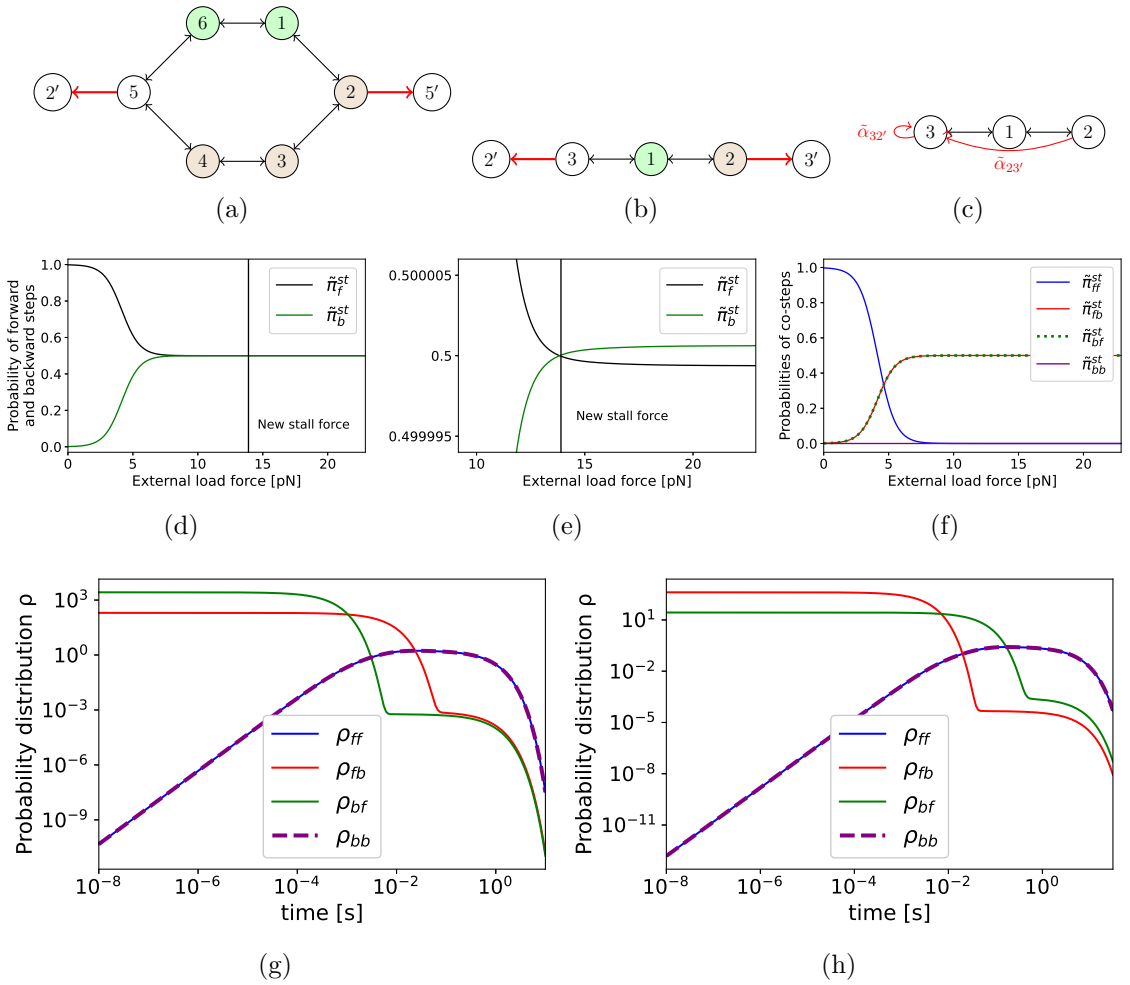


Figure B.5.: The coloured states in the 6-state model in (a) are merged. The coarse-grained system for the kinesin motor is depicted in (b). Again, the absorbing mechanical transition rates are coloured in red. The closed diagram (c) allows the calculation of the absorption flux according to Hill's method [7] if the system starts in state 3 after a forward step. In (d) and (e), the probabilities of forward and backward steps are depicted. In (f), the probabilities of co-steps in the coarse-grained system are plotted against the external load force. The co-steps distributions are depicted for zero load force in (g) and for $F = 5\,\mathrm{pN}$ in (h).

where the primed probabilities refer to the steady-state probabilities in the closed system in figure B.5(c). The absorbing flux for initial state 2, which is occupied after a previous backward step, can be calculated as described in the previous sections. Comparing figures B.3(d) and B.5(d), we note that the stall force is changed. In the original tri-cyclic and in the coarse-grained tri-cyclic system, the steady-state probabilities π_f and π_b are equal for $F_{Stall} = 7\,\mathrm{N}$. Figure B.5(e) shows that the motor

has no net displacement for $\tilde{F}_{Stall} = 13.9\,\mathrm{pN}$. Reducing the number of fundamental cycle changes the stall force. Furthermore, the probabilities of co-steps are not retained either, cf. figures B.3(e) and B.5(f). For forces larger than the new stall force, the velocity of the motor is negative. The motor walks in backward direction slower than in the 6-state model as we have seen in figure 4.2 too. Alternating co-steps (fb and bf), that yield a zero net-displacement, are most likely. With increasing load forces co-steps ff and bb become unlikely. For load forces below the original stall force $F_{stall} = 7\,\mathrm{pN}$, the uni-cyclic network is a good approximation for the probabilities of the co-steps. Whereas the approximation error increases in the range of large load forces where the backward cycle is dominant. The backward cycle in the tri-cyclic 6-state model is lost by merging the green states in figure B.5(a). Therefore the 3-state model is suitable to describe the motor when the forward cycle is dominant and unsuitable when the deleted backward cycle is dominant.
The dwell time distributions of the co-steps in the uni-cyclic coarse-grained 3-state model can be computed according to equation (B.14) and are given by

$$\begin{aligned}
\tilde{\rho}_{ff}(t) &= \frac{\dot{\tilde{P}}_{33'}(t)}{\tilde{P}^{st}_{33'}} = \frac{\tilde{P}_{32}(t)\tilde{\alpha}_{23'}}{\tilde{M}_{ff}} \\
\tilde{\rho}_{bf}(t) &= \frac{\dot{\tilde{P}}_{23'}(t)}{\tilde{P}^{st}_{23'}} = \frac{\tilde{P}_{22}(t)\tilde{\alpha}_{23'}}{\tilde{M}_{bf}} \\
\tilde{\rho}_{fb}(t) &= \frac{\dot{\tilde{P}}_{32'}(t)}{\tilde{P}^{st}_{32'}} = \frac{\tilde{P}_{33}(t)\tilde{\alpha}_{32'}}{\tilde{M}_{fb}} \\
\tilde{\rho}_{bb}(t) &= \frac{\dot{\tilde{P}}_{22'}(t)}{\tilde{P}^{st}_{22'}} = \frac{\tilde{P}_{23}(t)\tilde{\alpha}_{32'}}{\tilde{M}_{bb}}.
\end{aligned} \tag{B.35}$$

Figures B.5(g) and (h) show the dwell time distributions for the co-steps in the coarse-grained uni-cyclic system for two values for the external load force. Coarse-graining from a 6-state model to a uni-cyclic 3-state model does not preserve the symmetry in the transition rates in the original model.

B.4.5. Comparison of dwell time distributions for different motor models

The dwell time distributions ρ_ν, $\nu \in \{ff, fb, bf, bb\}$ can be obtained from equations (B.27), (B.33) and (B.35) for the original and the coarse-grained models respectively. The time derivatives are obtained from the master equation which can be solved with

the propagator as described in equation (B.17). All four dwell time distributions for the co-steps are nonexponential as shown in figures B.6 and B.7, both for the original 6-state model and for the coarse-grained models. The non-exponential distributons demonstrate the non-Markovian character of the mechanical stepping process [22]. The dwell time distributions $\rho_{ff}(t)$ and $\rho_{bb}(t)$ for forward-after-forward and backward-after-backward steps decrease to zero for small dwell times and are identical in the original model. The dwell time distributions for forward-after-forward and backward-after-backward steps are identical due to symmetry in the transition rates in the original model. The symmetry is preserved in the 4-state system by merging two states in the forward and in the backward cycle respectively. The 3-state system does not exhibit symmetries in its transitions rates. The dwell dis-

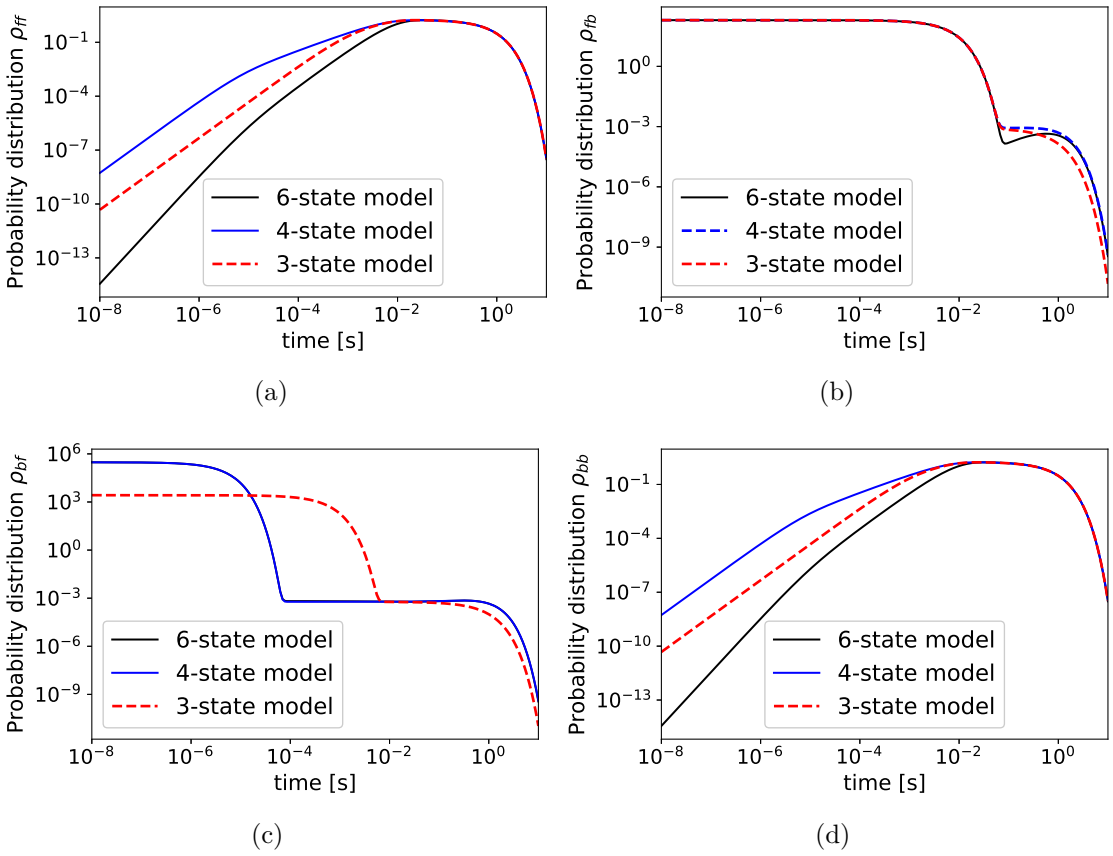


Figure B.6.: The dwell time distributions for the co-steps are computed for no external load force.

tributions ρ_{fb} and ρ_{bf} for forward-after-backward and for backward-after-forward steps exhibit several characteristic time scales. Both have a boundary maximum at

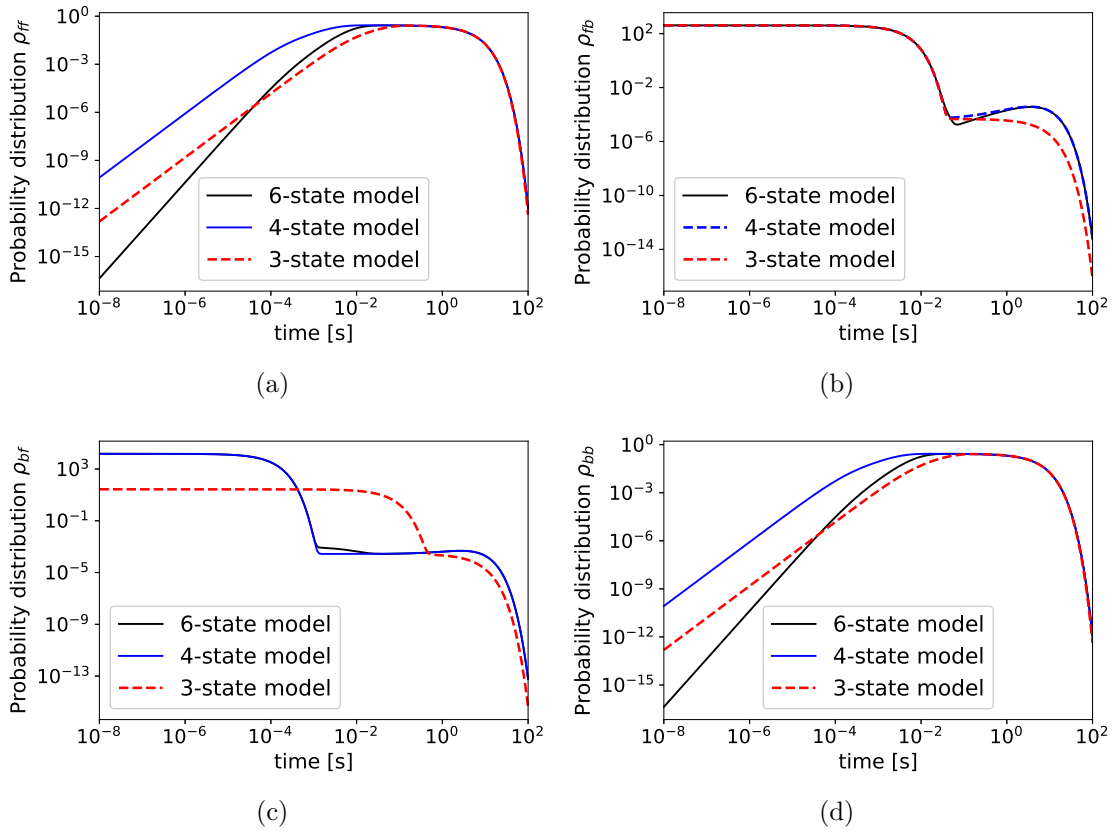
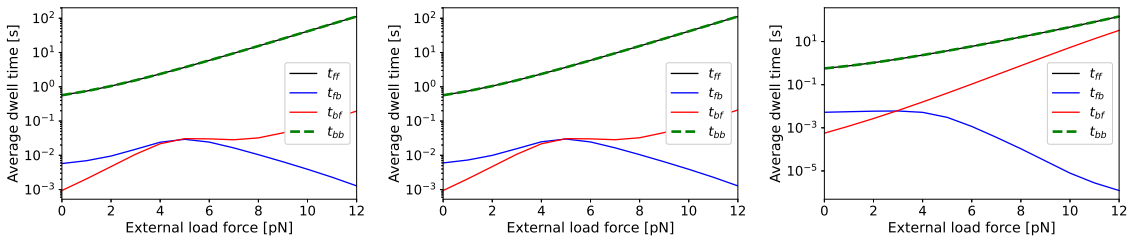


Figure B.7.: The dwell time distributions for the co-steps are computed for the load force $F = 5\,\mathrm{pN}$.

vanishing dwell times corresponding to two subsequent mechanical transitions without any intervening chemical transition [22]. In the original model, they exhibit a second shoulder at larger dwell times that correspond to cycles of intervening states chemical transitions that start from state 2 or state 5 and return to the same state before making the next mechanical transition [22]. For the coarse-grained uni-cyclic 3-state model, the distribution ρ_{bf} is shifted, cf. figures B.6(c) and B.7(c).

Figure B.8 shows the average dwell time for the co-steps as a function of the load force for different motor models. The distributions ρ_{ff} and ρ_{bb} are identical and hence the average dwell time for ff and bb co-steps is identical too. The backward cycle dominates for large load forces. Therefore the average of distributions ρ_{ff} and ρ_{bf} increases with increasing load because making a forward step is less likely. The average dwell times in the coarse-grained tri-cyclic 4-state model is not retained but is characterised by the same trends as the original model. The stall force is not retained in the uni-cyclic 3-state model and the average dwell times for the co-steps



(a) (b) (c)

Figure B.8.: Average dwell times for co-step distributions as a function of different kinesin motor models. In (a) for the original 6-state model. In (b) for the tri-cyclic 4-state model and in (c) for the uni-cyclic 3-state model.

are not retained either.

B.4.6. Summary

We can conclude that the tri-cyclic 4-state model approximates both the steady-state probabilities of steps and co-steps very well for a wide range of load forces. Both models are characterised by the same stall force. The uni-cyclic 3-state model resembling the forward cycle is a good approximation when the forward cycle is dominant, i.e. for zero or small load forces. The backward cycle becomes dominant for large stall forces and hence the uni-cyclic model depicted in figure B.5 is not a well-suited approximation anymore. As a matter of fact, the stall force is changed in the uni-cyclic model.

Both coarse-grained models have different dwell time distributions for the co-steps than the original model. The overall non-Markovian character is retained, but the distributions in coarse-grained models exhibit pronounced differences for small dwell times, cf. figures B.6 and B.7. The dwell time distributions for forward-after-forward and backward-after-backward steps are identical due to symmetry in the transition rates in the original model. The symmetry is preserved by the tri-cyclic model and hence the distributions for two steps of the same kind are still identical. The symmetry is broken in the uni-cyclic system and hence $\tilde{\rho}_{ff}$ and $\tilde{\rho}_b b$ are still similar but not identical anymore. Coarse graining from 6 to 4 states and preserving the fundamental cycle structure does not change the overall trends of the average dwell times. Whereas coarse graining to a 3-state uni-cyclic system affects the average dwell times for the co-steps as depicted in figure B.8.

Our coarse-grain approach was designed to preserve the net numbers of transitions

between any pair of state and hence our approach involves approximation errors for the dynamic. We can state nevertheless that preserving the network topology of a system yields to a smaller approximation error for dwell time distributions.

C. Supplementary Material: Parametrisation of the chemotaxis adaptation network

The parameters for the free energy in the Monod-Wyman-Changeux like model are from [26] and are used to coarse grain the chemotaxis adaptation network in section 4.2. The chemosensory system of E. coli and its environment can be characterized by three variables. Firstly the ligand concentration s (the input), secondly the average kinase activity a (the output) and thirdly the average methylation level m of the receptors (the memory) [23]. The receptors are distributed all over the membrane and interact with nutrient molecules in the environment. Binding of attractant ligands inhibits the kinase activity [24, 26]. The system adapts due to methylation or demethylation of receptors which is controlled by the receptor activity [25]. Binding of a ligand changes the kinase activity of chemotaxis protein CheA that phosphorylates a messenger protein CheY. The phosphorylated CheY induces tumbling. If receptors bind attractants, the rate of CheY phosphorylation is reduced. Therefore the tumbling frequency decreases and the bacterium moves up the gradient. The methylation level of the chemoreceptors allows adaptation. Methylation enhances the kinase activity and is responsible for adaptation to changes in ligand concentration [23, 25].

The free energy difference $\Delta E(s, m)$ between active and inactive states depends on the ligand concentration s (the signal) and the methylation level m by

$$\Delta E(s, m) = E_1(m) - E_0(m) = N e_m(m_1 - m) + N \ln\left(\frac{1 + s/K_I}{1 + s/K_A}\right), \qquad \text{(C.1)}$$

where $N = 1$ is the number of strong coupled dimers, $e_m = 2$, $m_1 = 1$, the dissociation constant $K_I = 18.2\mu M$ for the inactive receptors and $K_A = 3000\mu M$ for the active receptors. The activation rate $\omega_0(m)$ (corresponding to $\alpha_{i+1,i}$ in figure 4.3)

and the deactivation rate $\omega_1(m)$ (corresponding to $\alpha_{i,i+1}$) depend on the free energy difference of the active and inactive states [24]. The methylation level m and state i in figure 4.3 are connected by $m = \lfloor i/2 \rfloor$. The proportion of deactivation and activation rate is given by

$$\frac{\omega_1(m)}{\omega_0(m)} = \frac{\alpha_{i,i+1}}{\alpha_{i+1,i}} = e^{\Delta E(s,m)} = e^{E_1(s,m)-E_0(s,m)}. \tag{C.2}$$

The methylation rate $k_{+,a}(m) = \alpha^a_{i,i+2}$ and demethylation rate $k_{-,a}(m) = \alpha^a_{i+2,i}$ for inactive receptors ($a = 0$) or active receptors ($a = 1$) have the proportion

$$\begin{aligned} \frac{k_{+,0}(m)}{k_{-,0}(m)} &= \frac{1}{\gamma_0} e^{Ne_m/2} \\ \frac{k_{+,1}(m)}{k_{-,1}(m)} &= \gamma_1 e^{-Ne_m/2}. \end{aligned} \tag{C.3}$$

The methylation rate for inactive receptors is set to be $\alpha^{a=0}_{i,i+2} = k_{+,0}(m) = 1s^{-1}$. The factor γ_0 and γ_1 (later $\gamma_0 = \gamma_1 = \gamma$ for simplicity) break detailed balance. For $\gamma_0 = \gamma_1 = 1$ the system is in equilibrium. The parameter γ_0 suppresses the demethylation rate for inactive receptors from its equilibrium value, while γ_1 suppresses the methylation rate for active receptors. The thermodynamic force driving the cycle $(i, i+2, i+3, i+1, i)$ (see figure 4.3 with methylation $m = \lfloor i/2 \rfloor$) is given by

$$\Delta S = \ln\left(\frac{\omega_0(m+1) \cdot k_{-,1} \cdot \omega_1(m) \cdot k_{+,0}}{\omega_1(m+1) \cdot k_{+,1} \cdot \omega_0(m) \cdot k_{-,0}}\right) = \ln\left(\frac{1}{\gamma_0\gamma_1}\right) \tag{C.4}$$

The thermodynamic force is zero, if $\gamma_0 = \gamma_1 = 1$.

Danksagung

An dieser Stelle möchte ich all jenen danken, die durch ihre fachliche und persönliche Unterstützung zum Gelingen dieser Masterarbeit beigetragen haben. Zuerst gebührt mein Dank Herrn Professor Klumpp für die intensive Betreuung und die inspirierenden Besprechungen. Außerdem bedanke ich mich bei Herrn Professor Sollich für die Literaturempfehlungen und kritischen Anmerkungen zu meinem Manuskript. Der Studienstiftung des deutschen Volkes möchte ich für die ideelle und finanzielle Unterstützung danken, die ich während meines gesamten Studiums der Physik genießen durfte. Vielen Dank an die aufmerksamen Korrekturleser Shayan Hundrieser und Cai Dieball. Und zuletzt vielen Dank an meine Eltern und Geschwister, die mir eine angenehme Studienumgebung in Coronazeiten bereitet haben.

Erklärung Ich versichere hiermit, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig verfasst und nur die von mir angegebenen Quellen und Hilfsmittel verwendet habe. Wörtlich oder sinngemäß aus anderen Werken entnommene Stellen habe ich unter Angabe der Quellen kenntlich gemacht. Die Richtlinien zur Sicherung der guten wissenschaftlichen Praxis an der Universität Göttingen wurden von mir beachtet. Eine gegebenenfalls eingereichte digitale Version stimmt mit der schriftlichen Fassung überein. Mir ist bewusst, dass bei Verstoß gegen diese Grundsätze die Prüfung mit nicht bestanden bewertet wird.

Göttingen, den September 1, 2020

(David Seiferth)