Entropy, Dissipation and Information in Models of Interacting and Coarse Grained Systems

Maximilian Eisbach

Faculty for Mathematics and Natural Science
Institute of Theoretical Physics
TU Berlin

Supervisor
Prof. Dr. Tobias Brandes
Co-Supervisor
Prof. Dr. Alexander S. Mikhailov

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Physics

Berlin 2017

Acknowledgements

I am indebted to Prof. Tobias Brandes, Prof. Harald Engel and Prof. Alexander Mikhailov for their sincere support during the course of my studies. They have been very inspiring people for me who guided me all along the way. Both in academical as well as personal regards.

Much of my learning was guided by Philipp Strassberg, Gernot Schaller and Javier Cerrillo Moreno. They were great co-workers at the TU Berlin. Especially I want to thank Dimitri Loutchko for working with me at the Fritz-Haber-Institut.

A lot of my joy in theoretical physics originated from the books and lectures by Prof. Leonard Susskind from Stanford University. I am dignified to his way of teaching physics. Crystal clear, inspirational and entertaining. I have put little exceprts from his book "Black Hole Wars", which has been a real joy to read.

It is also my strong wish to acknowlegde the DFG (German research foundation) for financial support within the Collaborative Research Center 910 (SFB 910).

Abstract

This thesis studies information theoretic approaches for the motivation and application of the formula of the entropy production rate of Markov processes as well as the analysis of interacting stochastic processes.
As a particular example, the stochastic process describing the cathalytic cycle of tryptophan synthase is investigated in the framework of stochastic thermodynamics. The enzyme tryptophan synthase is characterized by a complex pattern of allosteric interactions that regulate the catalytic activity of its two subunits and opening or closing of their ligand gates. As a single macromolecule, it implements 13 different reaction steps, with an intermediate product directly channeled from one subunit to another. Here, a model based on experimental data is used to consider stochastic thermodynamics of such a chemical nanomachine. The Gibbs energy landscape of the internal molecular states is determined, the production of entropy and its flow within the enzyme are analyzed and the information exchange between the subunits resulting from allosteric cross-regulations and channeling is discussed.
Research from the 80s on the critical dynamics of spin systems described by a master equation is reviewd and put into context of recent developments from the field of stochastic thermodynamics. A method for the Renormalizration Group transformation of the master equation of a single, non-interacting, blockspin is introduced. This method allows the straight forward recursive RG transformation of a blockspin.

Contents

Part I

Introduction & core concepts

Introduction

Non-equilibrium Statistical Mechanics is a very broad field of studies. Possibly all physical systems are not truly isolated systems. Dissipation may arise due to reasons such as unknown fundamental theories, measurement processes, emergent phenomena from complexity or environmental noise. Equilibrium Statistical Mechanics describes the special type of systems which are thermalised, i.e. systems which have already relaxed towards an equilibrium state. Under equilibrium conditions, no energy, no matter or more generally no entropy is exchanged between the system and the environment. In turn systems undergoing thermalisation, i.e. relaxation towards a steady state or systems which have already relaxed towards a steady state and display constant entropy flows towards the environment belong to the huge class of non-equilibrium processes.
Above and beyond, a central question is at which scale in space and time a physical system only allows a statistical description. Statistical Mechanics is currently challenged by the nano-sciences which provide many examples of systems, often out of equilibrium, of size between atoms and macro systems. Thus it may be difficult to draw a line between the mircroscopic behaviour where deterministic atomic features are apparent and the macroscopic behavoir, where the large number of particles, calls for a statistical description. The new perspective provided by nano-sciences endorses a the development of non-equilibrium Statistical Mechanics in order to explain properties of these intermediate systems. Indeed nano-systems may be small enough to be considered as dynamical systems but also large enough to for the onset of statistical behavoir.

This thesis is outlined as follows: In chapter 1 the meaning of information conservation in the fundamental physical laws is reviewd. Coarse grained deterministic dynamical systems generally follow probablistic time evolution laws which, on average, may not

obey time reversal symmetry. For these system statistical time reveral symmetry can be broken and the accessible information about the state of the system changes at different pace in the future and past. In chapter 2 Shannons entropy is introduced based on a combinatorial quantification of path histories of a system ensemble. With the concept of entropy rates the breaking of statistical time reversibility is quantified and a connection to the thermodynamic entropy production and the second law is established. Interacting systems allow for an even richer information theoretic analysis as presented in chapter 3. All systems investigated in this thesis are described by a so called master equation, obeying the Marov assumption and describing the time evolution of the probability distribution. The key mathematical properties of the continuous time master equation are introduced in chapter 4. In chapter 5 the concept of information flow between sub-systems performing mutual feedback on each others but undergoing seperate fluctuations is introduced. This theory is expanded to systems which are in parts directly coupled to the same fluctuations. The chemical nano-machine Tryptophan synthase is analysed in chapter 6 in the framework of Stochastic Thermodynamics and the influence of the feedback- and direct-counpling is investigated. Chapter 7 is a summary of the authors latest but unfinished research in the field of critical slowing down of spin systems driven close to their critical point. The aim is to investigate the Kibble Zureck mechanism, i.e. the slow cooling of a spin system towards its critical point, where the system will fall into a non-equilibrium state while thermodynamic equilibrium constraints are imposed by the environment.

Chapter 1

Information in physical systems at different times and scale

> "Out of every 10,000,000,000 bits of information in the universe, 9,999,999,999 are associated with the horizons of black holes. It should be evident that our naive ideas about space, time, and information are wholly inadequate to understand most of nature."
>
> *Leonard Susskind* in *Black Hole Wars*

1.1 Information conservation in the fundamental physical laws

One might expect that a fact as basic as the existence of time's arrow would be embedded in the fundamental laws of physics. But the opposite is true. The fundamental equations of Quantum Mechanics do not distinguish between the past and the future: Any pair of distinct Quantum states remains distinct in the future and past. This fact is closely related to the conservation of information in the future and past [59]. Since we will formalize the notion of information later, we start with a heuristic definiton in order to grasp the connection between information conservation and the fundamental physical laws [53]: *"Information is the ability to distinguish reliably between possible alternatives."* In the language of Quantum Theory 'possible alternatives' are orthogonal states. Orthogonal states from the past are orthogonal now and

remain orthogonal in the future for a closed isolated system [57].

$$\langle j(0)|\, U^{\dagger}(\Delta t)U(\Delta t)\,|i(0)\rangle = \delta_{ij} \qquad \forall \Delta t \tag{1.1}$$

Unitary time evolution of state vectors in Quantum Mechanics is the most fundamental law in physics, we use to describe nature. It posses time inversion symmetry (since valid for all Δt) and also preserves any distinctions (orthogonality) between states - in other words information - as they propagate in time.
Figuratively speaking, the equation for the time evolution of a system dictates the neighboring structure of prior and posteriour states in time. It is a set of rules determining all possible states a given state will transform into after an infinitessimal time dt. These rules can be formulated in terms of a network structure connecting states as time evolves.
If this network connects states at time $t = 0$ only with one prior at $t = -dt$ and posterior state at $t = +dt$, then there is no ambiguity into which state the system can evolve and has come from at a time t. Any state has a unique past and future. Systems following such evolution laws are called *conservative determinsitic dynamical systems* [58, 57].
On the other hand, once there is more than one possibility for a past and future of a state, everything suddenly becomes more challanging. Systems of this class are *dissipative determinsitic dynamical systems* and *stochastic dynamical systems*. It is much harder, genrally impossible, to predict the future and reproduce the past in both cases since they may obey highly complicated dynamics [4, 44].
While the conservation of distinctions is held so sacred in the fundamental laws, it is astonishing to witness that this perspective is practically inaccessible for the analysis of most interesting systems. Inevitably, violation of information conservation must be considered to play a fundamental role in our description of the physical world.

1.2 Coarse graining of states space and time

As opposed to equation (1.1) broken time reversal symmetry appears whenever our current information about a system does not allow us to exactly predict its future state and reproduce its past for all times. Time reversal symmetry is broken, when our information about a system tends to loose its precision with different speed into future and past as time unfolds. This is often due to the fact that we cannot describe

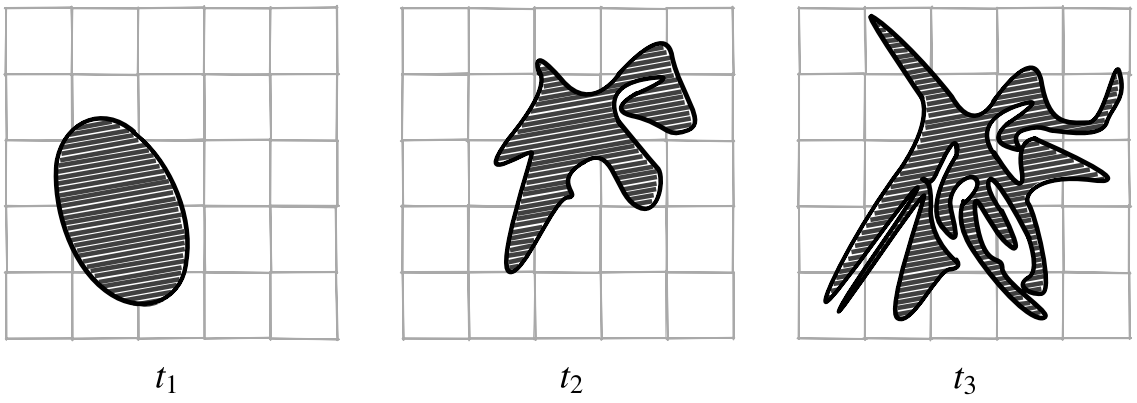

Figure 1.1: **(1-3)** The fractal like spreading of deterministic trajectories of a dynamical system can lead to an unpredictable behavior of the system on a coarse grained level. The rapid spreading of the trajectories and the resulting unpredictable paths through the phase space cells is commonly referred to as the dynamical randomness of a system.

elementary processes precisely enough: In laboratory experiments our measurements will always have errors and a lower resolution in space and time than all the processes happening on the scales of the fundamental laws. On the pure mathematical side, we are simply unable to exactly solve the underlying equations in many cases [56, 23]. For both, the experimental and mathematical limitations, a commonly successful approach is to ignore processes at scales we anticipate to be irrelevant for the phenomena we whould like to understand. Following such a scheme of approximation is generaly called *coarse graining*. Degrees of freedom at irrelevant scales are lumped together to fewer degrees of freedom. The incentive is to simplify the problem [12]. In other words: We are sloppy with the little things happening in complicated systems in order to frame a problem into models reflecting the phenomena of interest which we can explain and analyse.

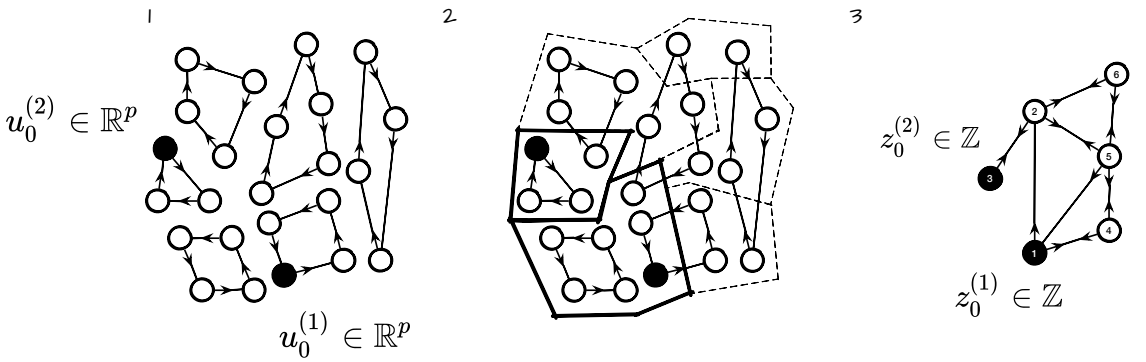

Figure 1.2: **(1)** In the illustration, states can only evolve into states connected by a single link. For systems governed by the fundamental laws of physics, the time evolution of two distinct initial states $u_0^{(1)}$ and $u_0^{(2)}$ (marked black) will never fall onto one and the same state. In that sense information is conserved. **(2)** Our limited acces to the microscopic degrees of freedom divides the state space in into cells containing all configurations which represent a single state in a coarse grained model. **(3)** The mapping of the states onto cells can lead to a time evolution of transitions between sectors which is not time inversion symmetric. These dynamics of a coarse grained model do not necessarily preserve information (see text).

Coarse graining deterministic systems

Now, we exploit the possible consequences to the time evolution of a conservative system by analysing the dynamics of a coarse grained state space

$$d_t \boldsymbol{u}(t) = \boldsymbol{f}(\boldsymbol{u}(t)), \qquad \mathbb{R}^p \mapsto \mathbb{R}^p \qquad \textit{classical deterministic dynamical system on a continous state space evolving in continous time} \tag{1.2}$$

$$\sigma(t) = \boldsymbol{g}(\boldsymbol{u}(t)), \qquad \mathbb{R}^p \mapsto \mathbb{N} \qquad \textit{a surejective mapping which represents the limited resolution of the measurment space} \tag{1.3}$$

where $\boldsymbol{f}(\bullet)$ is a bijective mapping and $\boldsymbol{g}(\bullet)$ is a surjective mapping. The role of the surjective mapping is the coarse graining of the underlying state space into cells, which can be labeled by integer numbers, i.e. $\sigma(t) \in \{1, 2, ..., N\}$ (see figure 1.2). Since a surjective mapping is not invertible it is clear that it is a source of information loss. We call

$$\boldsymbol{\sigma}_t \coloneqq (\sigma_0, \sigma_1, ..., \sigma_t)^\intercal, \tag{1.4}$$

a typical path or sequence of a realization of the stochastic process $\boldsymbol{\sigma}_t$ further on.

1.3 Dynamical Randomness

A coarse grained state $\sigma(t)$ can be followed by a variety of different outcomes. The actual sequence ultimately depends on the initial configuration $\boldsymbol{u}(0)$ of the deterministic dynamical system.

As a consequence of the surjective mapping (1.3), a system described on the level of a sequence of coarse grained states evolves in time due to a rich possibility of all trajectories of $\boldsymbol{u}(t)$ traversing different cells (see figure 1.2). The system $\boldsymbol{u}(t)$ may trace complicated, fractal like trajetories in its state space [44] and therefore traverse many different sectors (see figure 1.1). We will see below that this ambiguity leads to an observation of the time evolution of paths $\boldsymbol{\sigma}_t$ which may obey *statistical temporal asymmetry*, in the sense that not given the initial condition $\boldsymbol{u}(0)$ we cannot exactly determine the future outcomes $\boldsymbol{\sigma}_{t'>t}$ or reproduce their past sequence from the current knowledge of $\boldsymbol{\sigma}_t$ alone.

The extensive neighboring structure of the time evolution network for the coarse grained states with more than one ingoing and outgoing connections to other coarse grained states gives inherently random, logically irreversible sequences of states. Such dynamics are called *stochastic processes* or *random walks*, since without the exact knowledge of all underlying dynamics, they seem to randomly evolve in time.

As opposed to solutions of an ordinary differential equation a stochastic process can evolve in many, sometimes infinitely many ways. Such a process is generically characterised by both a random variable for the sequence of states $\boldsymbol{\sigma}_t$ and one for the jump times $\boldsymbol{\tau}_t$ [65].

1.4 Statistical temporal asymmetry

Due to the uncertainty arising from the coarse graining procedure the information about the state of the system can only be given as a probability distribution $\mathcal{P}(\boldsymbol{\sigma}_t)$ of a symbolic sequence over some time interval $[0:t]$. We can investigate whether this probabalistic information of a coarse grained system is conserved in the future and past, i.e. whether the probability distributions are distinguishable. Whenever this information is conserved for all times in the past and future sequence of states, we

obviously have time reversal symmetry with regards to the probablistic description, i.e.

$$\mathcal{P}(\boldsymbol{\sigma}_{-t}) = \mathcal{P}(\boldsymbol{\sigma}_t). \tag{1.5}$$

If the probability distribution of future and past sequences aren't symmetric we say that information is not conserved and

$$\mathcal{P}(\boldsymbol{\sigma}_{-t}) \neq \mathcal{P}(\boldsymbol{\sigma}_t). \tag{1.6}$$

In translating the time evolution of systems into sequences of measurment outcomes, better yet the probability of these sequences, we have arrived at the core conceptual starting points of the thesis: We are interested in the transformation of the reversibility of the time evolution for systems at different levels of time resolution and scale. We take the standpoint that temporal asymmetry in physical systems may naturally emerge from our own ignorance and ability to follow all processes in detail. Temporal asymmetry of coarse grained processes is studied in terms of time series of stochastic processes and their associated probability distributions.

1.5 Markov and stationary processes

In general the sequence of the symbols in $\boldsymbol{\sigma}_t$ is not drawn independent and identically distributed random variables. Interactions within the system and the environment lead to stochastic processes where the sequence of symbols $\boldsymbol{\sigma}_t$ is not arbitrary. A simple example of such a stochastic process with dependence among the symbols is one in which each random variable $\sigma \in \boldsymbol{\sigma}_t$ is only dependent on the one preciding it. Especially it is conditionally independent of all the other preciding random variables from the past. Such a type of process is called a *Markov process*, where

$$p(\sigma_{t+dt}|\sigma_t) = p(\sigma_{t+dt}|\boldsymbol{\sigma}_t) \tag{1.7}$$

is commonly called the Markov property. Under this condition the probability of a particular realisation of the process becomes

$$p(\boldsymbol{\sigma}_t) = p(\sigma_0)p(\sigma_1|\sigma_0)p(\sigma_2|\sigma_1)...p(\sigma_{t-dt}|\sigma_t). \tag{1.8}$$

A stochastic process is said to be *time invariant* or *stationary* if the conditional probability $p(\sigma_{t+dt}|\sigma_t)$ does not depend on t, i.e.

$$p(\sigma_{t+dt}|\sigma_t) = p(\sigma_{t+\Delta t+dt}|\sigma_{t+\Delta t}). \qquad \forall \Delta t \tag{1.9}$$

Chapter 2

Entropy rates and statistical temporal asymmetry

> "The entropy of a car and the entropy of a rust heap have something to do with the number of arrangements that we would recognize as rust heaps versus the number that we would recognize as a car."
>
> *Leonard Susskind* Black Hole Wars

2.1 Information content as the average code length

Physicists usually gather data about a system and deduce a hopefully simple model, from which one can predict and understand the time evolution of the system. As a fact, the success of building a model may highly depend on how well we can classify the data points and detect patterns to draw the right conclusions about the dynamics and interactions of the system. *Information Theory* gives us the tools to quantify the notion of information [14] and randomness in data generated when observing a system. In order to distinguish all possible states of a system, binary codes can be used. The sequence $\boldsymbol{\sigma}_t$ can simply be seen as an abstract message where the different states are signals or data points to be encoded into binary strings [15]. The tricky part is to actually find a suitable, compressed code that we can handle in practise since we aim to build simple laws predicting the content of these binary strings for the future and past. As we saw in the preceeding chapter, the sequence $\boldsymbol{\sigma}_t$ can have a huge variety of outcomes over time. It is the concept of *entropy* which allows us to

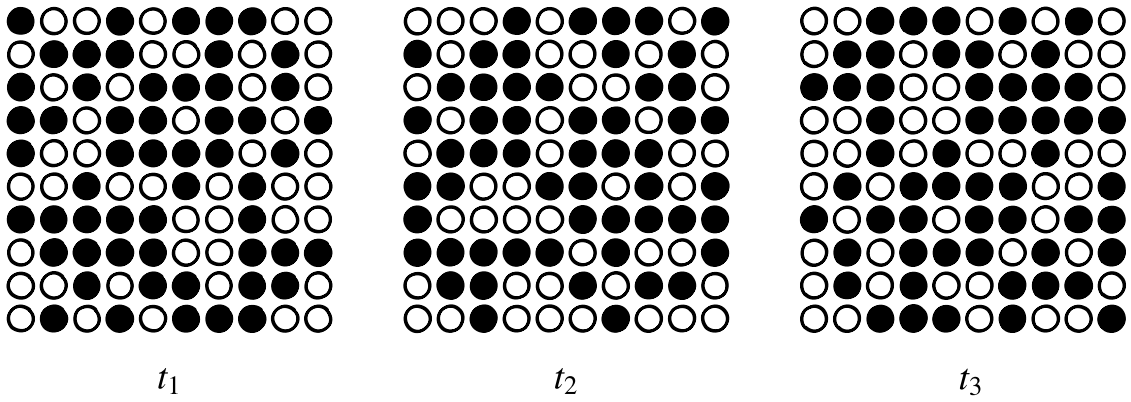

Figure 2.1: A decription of a system depends on the state of our knowledge. Here all three configurations look similar at different times. Clearly by visual inspection it is not possible to find a law which exactly describes the time evolution of the different configurations. We can e.g. count how often black appears on the cubes and try to deduce a law for the time evolution. This would result into a probabalistic discription. But I can also tell you that the configuration is rotated by 90 degrees from t_1 to t_2 and t_3, from which we can clearly deduce a deterministic description.

quantify the average number of bits needed to encode these time series of a system, which simultaneously reflects the variety of outcomes.

Estimating the number of arrangements that conform to a specific recognizable criterion

After all we want to quantify how many bits are needed to encode the sequences of the signal $\boldsymbol{\sigma}_t$ into binary strings. Of course this will vary from each realisation of the stochastic process, s.t. whe can only determine the statistical average of bits needed. So we better ask: How many different sequences $\boldsymbol{\sigma}_t$ are there in an ensemble

$$\{\boldsymbol{\sigma}_t\} \coloneqq \lim_{N\to\infty} \left\{\boldsymbol{\sigma}_t^{(1)}, \boldsymbol{\sigma}_t^{(2)}, ..., \boldsymbol{\sigma}_t^{(N)}\right\} \tag{2.1}$$

of N copies of the system, observed during a time interval $[0:t]$, and how likely are they? Taken literally, at time t we would classify all systems in the ensemble which have traced identical sequences $\boldsymbol{\sigma}_t$ of states and record the count by the number $N_{\boldsymbol{\sigma}_t}$. In the limit of very large N the probability that a single system traces the sequence $\boldsymbol{\sigma}_t$, can be defined as the frequency of occurences in the ensemble

$$p(\boldsymbol{\sigma}_t) \coloneqq \lim_{N\to\infty} \frac{N_{\boldsymbol{\sigma}_t}}{N}. \tag{2.2}$$

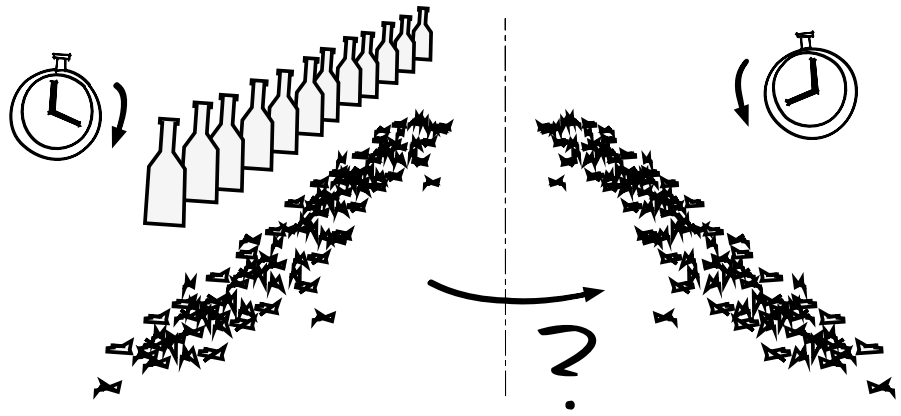

Figure 2.2: When we drop an ensemble of bottles, then there is almost an infinite number of ways the bottles will burst into pieces. Yet there is only a limited number of ways in how the pieces of the bottles can reassmble again such that we recognise them as the initial bottles.

Since the copies in the ensemble are statistically independent, the number of all possible realisations of the distribution $\mathcal{P}(\boldsymbol{\sigma}_t)$ is given by (*Boltzmann 1877* [46])

$$\mathbb{W}[\boldsymbol{\sigma}_t] \;\coloneqq\; \frac{N!}{\prod_{\{\boldsymbol{\sigma}_t\}} N_{\boldsymbol{\sigma}_t}!}. \tag{2.3}$$

Definition (2.3) counts the many ways in which the ensemble could evolve to produce a particular distribution. The larger (2.3), the more difficult it is to distinguish between all this possibilities. At the same time the number of bits needed to reliably record the time evolution grows. For large N the combinatorial number (2.3) is proportional to (see Appendix B.1),

$$\mathbb{W}[\boldsymbol{\sigma}_t] \;\sim\; \exp\Big(-\sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t)\log p(\boldsymbol{\sigma}_t)\Big) \tag{2.4}$$

$$= \exp\Big(\mathbb{H}[\boldsymbol{\sigma}_t]\Big), \tag{2.5}$$

Thus (2.3) *grows* exponentially with a rate called the *Shannon entropy* [14]

$$\mathbb{H}[\boldsymbol{x}] \coloneqq -\sum_{\boldsymbol{x}} p(x)\log p(x) \geq 0. \tag{2.6}$$

Note that the actual base of the logarithm is arbitrary and determines the units (bits, nats etc.) we measure information with (see Appendix B.2). The Shannon entropy

can be interpreted as an estimate of hidden, undetected information or uncertainty in a random variable $\boldsymbol{x}$ and its associated probability distribution $\mathcal{P}(\boldsymbol{x})$. With regards to coarse grained deterministic synamical systems it is a measure of our ignorance we employ when lumping the original dynamics (1.2) into coarse grained dynamics (1.3). The dynamics of a deterministic conservative dynamical system don't produce entropy since there is no ambiguity in the way how a probability distribution of such a system evolves. This fact is also known as *Liouville's theorem* [32]. Shannons entropy (2.6) is maximized if all random variables are equally likely and grows with the number of possible outcomes x. Hence, the hidden information about the state of the ensemble $\{\boldsymbol{\sigma}_t\}$ grows if the number of possible sequences grows in time. The growth rate is maximized when they are equally likely, i.e. indistinguishable [14].

Reversed paths

Based on a probablistic description of the system we can quantify how much the information about a system differs when time evolves into future and past. Certainly, every once in a while, the system will produce sequences of outcomes $\boldsymbol{\sigma}_t^R \coloneqq (\sigma_t, \sigma_{t-dt}, ..., \sigma_0)^\intercal$, which happen to be exactly the reverse sequence of $\boldsymbol{\sigma}_t \coloneqq (\sigma_0, \sigma_{dt}, ..., \sigma_t)^\intercal$, as if time would have been reversed. This is possible for any reversibel stochasitc process and lies at the heart of the requirement of *thermodynamic reversibility* in physics. This requirement can be employed by the use of an involution with the property $\boldsymbol{\sigma}_t^R = \hat{\Theta}\boldsymbol{\sigma}_t$ and $\boldsymbol{\sigma}_t = \hat{\Theta}^2\boldsymbol{\sigma}_t$. We demand that $\hat{\Theta}\boldsymbol{\sigma}_t \neq \emptyset$ for all $\boldsymbol{\sigma}_t$, which again is the requirement of thermodynamic reversibility, i.e. each sequence can be reversed. We now aim to quantify the ratio

$$\frac{p(\boldsymbol{\sigma}_t)}{p(\hat{\Theta}\boldsymbol{\sigma}_t)} \coloneqq \exp\Big[s(\boldsymbol{\sigma}_t)\Big]. \tag{2.7}$$

in terms of the exponent $s(\boldsymbol{\sigma}_t)$, in in order to estimate the statistical temporal asymmetry of future and past sequences. According to (1.9), if $s(\boldsymbol{\sigma}_t) = 0$ then we say information about the system is conserved in the future and past. On the other hand, when $s(\boldsymbol{\sigma}_t) \neq 0$ the information is not conserved, since the future and the past of a sequence are not equally likely.
For statistical time reversibility the following equality holds [48]

$$p(\boldsymbol{\sigma}_t) = p(\boldsymbol{\sigma}_t^R) \tag{2.8}$$

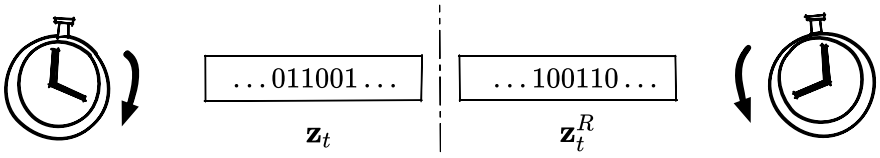

Figure 2.3: Sequences 0100011... and their reverse ...1100010 are denoted by $\boldsymbol{\sigma}_t$ and $\boldsymbol{\sigma}_t^R$.

which is equal to say that a sequence and its reverse are always equally likely when we, formally, let time evolve in opposite directions in both cases.
Note hat (2.8) demands the sublte but important fact that we only observe one single ensemble and therfore don't need to worry about matching the right initial conditions, which would be the case if we'd run a new, seperate experiment to investigate the time reversal [62]. Taking this restriction into account (2.7) can equally be written as

$$s(\boldsymbol{\sigma}_t) = \log \frac{p(\boldsymbol{\sigma}_t)}{p(\boldsymbol{\sigma}_t^R)}. \tag{2.9}$$

This means we can equally look at the time reversal of an ensamble by counting how often a sequence $\boldsymbol{\sigma}_t$ is actually realized and, for its time reversal, we simply count $\boldsymbol{\sigma}_t^R$ (see figure 2.4). We obtain $\boldsymbol{\sigma}_t$ and $\boldsymbol{\sigma}_t^R$ by running a single experiment, in other words considering a single ensemble.

2.2 Information content of typical and reversed paths

Suppose we also have access to the number of how often the reversed sequences $\{\boldsymbol{\sigma}_t^R\}$ are actually realised by the ensemble within the time interval $[0:t]$. If the process possesses time reversal symmetry within $[0:t]$, we would assume that $\left\{N_{\boldsymbol{\sigma}_t^R}\right\} = \{N_{\boldsymbol{\sigma}_t}\}$ with absolute certainty, i.e. $p(\boldsymbol{\sigma}_t) = p(\boldsymbol{\sigma}_t^R)$. Whenever $\left\{N_{\boldsymbol{\sigma}_t^R}\right\} \neq \{N_{\boldsymbol{\sigma}_t}\}$, at least one sequence of the process must have a statistically preferred direction into future or past, thus breaking statistical temporal symmetry. The probability to find a reversed sequence $\boldsymbol{\sigma}_t^R$ can be defined by[1]

$$p(\boldsymbol{\sigma}_t^R) \coloneqq \lim_{N\to\infty} \frac{N_{\boldsymbol{\sigma}_t^R}}{N}. \tag{2.10}$$

[1]Counting on ergodicity, $\sum_{\boldsymbol{\sigma}_t^R} N_{\boldsymbol{\sigma}_t^R} = N$ holds.

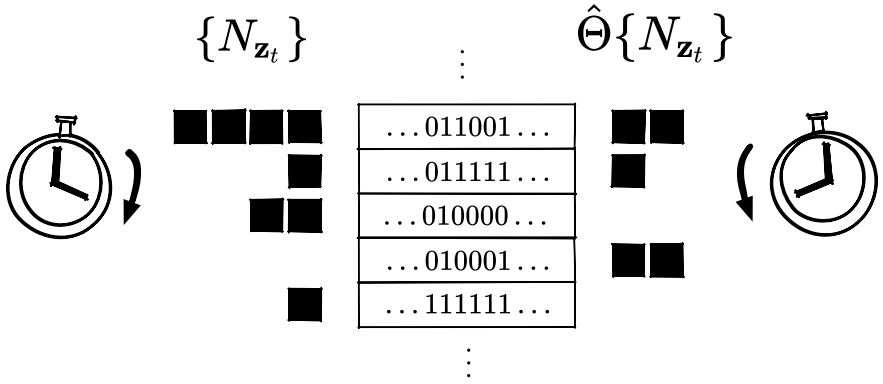

Figure 2.4: The N systems in the ensamble of stochastic processes is sorted according to the typical paths $\boldsymbol{\sigma}_t$ of the systems and their reverse $\boldsymbol{\sigma}_t^R$ (translated into binary strings). The numbers $N_{\boldsymbol{\sigma}_t}$ and $N_{\boldsymbol{\sigma}_t^R}$, counting how often the system traces the path $\boldsymbol{\sigma}_t$ and $\boldsymbol{\sigma}_t^R$, are going to be different whenever the ensemble dosen't trace path $\boldsymbol{\sigma}_t$ and $\boldsymbol{\sigma}_t^R$ equally often. In that sense the probability of observing a sequence in normal time and reversed time are different.

Hence, knowing all values of $N_{\boldsymbol{\sigma}_t^R}$ and $N_{\boldsymbol{\sigma}_t}$ can be seen as two different sequence classifiactions of the ensemble. To compare both information sources we are going to quantify how much uncertainty about the vaules $N_{\boldsymbol{\sigma}_t}$ remains once $N_{\boldsymbol{\sigma}_t^R}$ is known. In other words, we quantify how well the knowledge of the reversed sequences (the past if you will) suffices in order to encode the time evolution of $\boldsymbol{\sigma}_t$ with the same binary code. On the ensemble level this encodability condition is equal to write (see *Jaynes 2003* p.290 [31] or [46])

$$\mathbb{P}[\boldsymbol{\sigma}_t] \;\coloneqq\; \mathbb{W}[\boldsymbol{\sigma}_t] \prod_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t^R)^{N_{\boldsymbol{\sigma}_t}} \tag{2.11}$$

where $\mathbb{P}[\boldsymbol{\sigma}_t]$ is the likelyhood that a particular distribution $\{N_{\boldsymbol{\sigma}_t}\}$ is realized, given the prior knowledge of the probability distribution $\mathcal{P}(\boldsymbol{\sigma}_t^R)$ of the reversed sequences. Again with the use of Stirling's approximation for large N, this is (see Appendix B.1).

$$\begin{aligned}
\mathbb{P}[\boldsymbol{\sigma}_t] &\sim \exp\left(-\sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t) \log \frac{p(\boldsymbol{\sigma}_t)}{p(\boldsymbol{\sigma}_t^R)}\right) && (2.12)\\
&= \exp\Big(-\mathbb{D}_{\mathrm{KLD}}[\mathcal{P}(\boldsymbol{\sigma}_t) \parallel \mathcal{P}(\boldsymbol{\sigma}_t^R)]\Big) && (2.13)\\
&= \exp\Big(-\Big\langle s(\boldsymbol{\sigma}_t)\Big\rangle_{\{\boldsymbol{\sigma}_t\}}\Big) && (2.14)
\end{aligned}$$

where the brackets denote the ensemble average. Here, the combinatorial number (2.14) *decays* exponentially with an exponent called the *Kullback-Leibler Divergence* [14]

$$\mathbb{D}_{\mathrm{KLD}}[\mathcal{P}(\boldsymbol{x}) \parallel \mathcal{Q}(\boldsymbol{x})] \coloneqq \sum_{\boldsymbol{x}} p(x) \log \frac{p(x)}{q(x)} \geq 0. \tag{2.15}$$

which quantifies in bits how close a probability distribution $\mathcal{P}(\boldsymbol{x})$ is to a source (prior) distribution $\mathcal{Q}(\boldsymbol{x})$. The exponential decay is reasonable: The more copies of systems there are, the less likely that, just by chance, the number of occurances $\left\{N_{\boldsymbol{\sigma}_t^R}\right\}$ and $\{N_{\boldsymbol{\sigma}_t}\}$ are the same and thus $p(\boldsymbol{\sigma}_t) = p(\boldsymbol{\sigma}_t^R)$. With (2.14) we see that the ensemble average of (2.9) indeed quantifies the breaking of time reversal symmetry in terms of an information theoretical uncertainty between the probability distributions $\mathcal{P}(\boldsymbol{\sigma}_t)$ and $\mathcal{P}(\boldsymbol{\sigma}_t^R)$, namely

$$\Big\langle s(\boldsymbol{\sigma}_t) \Big\rangle_{\{\boldsymbol{\sigma}_t\}} = \mathbb{D}_{\mathrm{KLD}}[\mathcal{P}(\boldsymbol{\sigma}_t) \parallel \mathcal{P}(\boldsymbol{\sigma}_t^R)] \geq 0. \tag{2.16}$$

Equation (2.16) allows us to quantify the temporal asymetry or irreversibility of stochastic processes. If all sequences, forward and backward in time, are equally likely then (2.16) equals zero, and from observing $\boldsymbol{\sigma}_t$ we can predict the past with the same precision. If the probability of a time reversed sequences tends to zero, then (2.16) diverges.

Equation (2.16) gives rise to a very intuitive conclusion which can be derived for systems with statistical temporal asymmetry: The bigger (2.16), the more unlikely that the process can be described with equal probability distribution for its future and past for a time interval $[0 : t]$. From observing the forward process, we cannot deduce viable information about the reversed process. Therefore systems breaking statistical time reversal symmetry have the property of tracing sequences which are statistically very typical for the system in one direction but not for the reversed. We will see later that, the stronger time reversibility symmetry is broken, the better we can predict what sequences a system will trace in the future and yet the less can be reconstructed about its past.

2.3 Dynamical randomness and statistical temporal asymmetry as entropy production

As a next step we want to quantify how fast information about a system tends to loose its reliability when time evolves into future and past. This can be done by the concept of entropy per unit time, which has been introduced by Shannon [14] in the context of stochasitc processes and later by Kolmogorov [38] and Sinai [54] in dynamical systems theory. The entropy per unit time[23] is defined as

$$\dot{\mathbb{H}}[\boldsymbol{\sigma}] \;\coloneqq\; \lim_{t\to\infty} -\frac{1}{t}\Big\langle \log p(\boldsymbol{\sigma}_t)\Big\rangle_{\{\boldsymbol{\sigma}_t\}}. \tag{2.19}$$

This quantitiy is the average rate at which the amount of bits needed to record the sequence of the process $\boldsymbol{\sigma}_t$ changes when the system evolves dt further and refers to the dynamical randomness of a dynamical system. Similarly we can introduce an entropy rate, characterizing the disorder in time for the reversed paths

$$\dot{\mathbb{H}}^R[\boldsymbol{\sigma}] \;\coloneqq\; \lim_{t\to\infty} -\frac{1}{t}\Big\langle \log p(\boldsymbol{\sigma}_t^R)\Big\rangle_{\{\boldsymbol{\sigma}_t\}}. \tag{2.20}$$

The time-reversed entropy per unit time characterizes the dynamical randomness of the time-reversed paths in the forward process $\boldsymbol{\sigma}_t$.

[2]The supremum of the entropy rate (2.19) over all the possible coarse graining mappings $\boldsymbol{g}(\bullet)$ (c.f. equation (1.3)) defines the Kolmogorov-Sinai (KS) entropy per unit time of dynamical systems theory [13, 23].

$$\dot{\mathbb{H}}_{\mathrm{KS}} \coloneqq \sup_{\boldsymbol{g}} \Big(\dot{\mathbb{H}}[\boldsymbol{\sigma}]\Big) \tag{2.17}$$

Pesin's theorem states that the KS entropy is given by the sum of positive Lyapunov exponents for isolated dynamical systems [20]. For stochastic processes, the KS entropy depends on the sampling time dt as well as on the coarse-graining mapping $\boldsymbol{g}(\bullet)$. Randomness can be found at all spatial and time scales for stochastic processes, which calls for the entropy per unit time to diverge for small time and spatial resolution [24]. At the microscopic scale where the deterministic feature of the underlying dynamics (c.f. equation (1.2)) take over, this divergence saturates.

[3]In the continuous time limit, where $dt \to 0$, the entropy rate can be written as path integral expression, i.e.

$$\dot{\mathbb{H}}[\boldsymbol{\sigma}] \underset{dt\to 0}{=} \lim_{t\to\infty} -\frac{1}{t}\int \mathcal{D}[\boldsymbol{\sigma}_t]\mathcal{P}(\boldsymbol{\sigma}_t)\log\mathcal{P}(\boldsymbol{\sigma}_t), \tag{2.18}$$

where the path integral ranges over all possible stochastic trajectories.

With (2.19) and (2.20) we at the so called *entropy production rate* per unit time

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}] \coloneqq \dot{\mathbb{H}}^R[\boldsymbol{\sigma}] - \dot{\mathbb{H}}[\boldsymbol{\sigma}] = \lim_{t\to\infty} \frac{1}{t}\mathbb{D}_{\mathrm{KLD}}[p(\boldsymbol{\sigma}_t) \parallel p(\boldsymbol{\sigma}_t^R)] \geq 0, \tag{2.21}$$

which can also be formulated as a time dependent rate,

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}]_t \coloneqq \frac{\mathrm{d}\mathbb{D}_{\mathrm{KLD}}[p(\boldsymbol{\sigma}_t) \parallel p(\boldsymbol{\sigma}_t^R)]}{\mathrm{d}t} \geq 0 \tag{2.22}$$

which quantifies how much, on average, the probability distribution of the forward and backward process differ when the process evolves dt further. Under non-equilibrium conditions, i.e. $\mathcal{P}(\boldsymbol{\sigma}_t) \neq \mathcal{P}(\boldsymbol{\sigma}_t^R)$, the probabilities of the forward sqeuances and those of the corresponding time-reversed sequences break the time-reversal symmetry. This manifests itself in the fact that the decay rate (2.19) of the time-reversed paths is larger than the decay rate (2.19) of the typical paths. Their difference (2.21) is positive and gives the well-known thermodynamic entropy production for systems described by a master equation [22]. Thus the entropy production is directly related to the breaking of the time-reversal symmetry in the dynamical randomness of the nonequilibrium fluctuations. We note that under Markovian NESS coditons (see chapter 4.2.8)

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}] = \lim_{t\to\infty} \dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}]_t = \lim_{t\to\infty} \frac{1}{t}\Big\langle s(\boldsymbol{\sigma}_t)\Big\rangle_{\{\boldsymbol{\sigma}_t\}} \tag{2.23}$$

2.4 Predictability of stochastic processes

According to the *Shannon-McMillan-Breiman theorem* [8, 22], the path-probability indeed decays as

$$p(\boldsymbol{\sigma}_t) \sim \exp\left(-t\dot{\mathbb{H}}[\boldsymbol{\sigma}]\right), \tag{2.24}$$

for almost all the trajectories if the process is ergodic. This result shows that the more pronounced the spreading of the probabilities along all possible paths in the future, the higher will be the entropy per unit time. Consequently, the entropy per unit time can be seen as the dynamical randomness of the time evolution within $\boldsymbol{\sigma}_t$. Analogously for a typical path of the process, the probability to find the time-reversed path decays as

$$p(\boldsymbol{\sigma}_t^R) \sim \exp\left(-t\dot{\mathbb{H}}^R[\boldsymbol{\sigma}]\right). \tag{2.25}$$

Since the entropy production rate (2.21) is always positive, the exponential decay of (2.25) must be faster. In fact the ratio

$$\frac{p(\boldsymbol{\sigma}_t)}{p(\boldsymbol{\sigma}_t^R)} \sim \exp\left(\dot{\mathbb{H}}^R[\boldsymbol{\sigma}] - \dot{\mathbb{H}}[\boldsymbol{\sigma}]\right) = \exp\left(\dot{\mathcal{S}}_i[\boldsymbol{\sigma}]\right). \tag{2.26}$$

grows exponentially with a rate given by the entropy production rate. Therefore we see that systems under non-equilibrium conditions perform a selection of preferred paths when evolving into the future. The probability of typical pahts of the system decay more slowly than for the time-reversed paths. The process $\boldsymbol{\sigma}_t$ observed forward in time turns out to be more ordered than the backward process, since $\dot{\mathbb{H}}^R[\boldsymbol{\sigma}] \geq \dot{\mathbb{H}}[\boldsymbol{\sigma}]$ as imposed by the strictly positive entropy production. From a more information theoretic viewpoint, we can say that the entropy production measures the information cost of describing the temporal evolution in the time-reversed frame of reference. Nonequilibrium processes generate more dynamical order or information than equilibirum processes. This is an interesting perspective on the second law of thermodynamics: Nonequilibrium conditions perform a selection of preferred paths and time ordering of the trajectories occur as soon as the system is driven out of equilibrium. Spin-systems such as the *Ising-Model* are excellent examples to study the connection of information theoretic quantities to thermodynamic quantities such as the free energy when approaching the critical point. The author believes that the above introduced theory is a worthwhile new direction to look at the *Kibble-Zurek-Mechanism* and the critical slowing down.

Chapter 3

Entropy rates and interacting systems

"In this John Wheeler world of information, the laws of physics would consist of rules for how the configuration of bits is updated from instant to instant. Such rules, if correctly constructed, would allow waves of Os and Xs to propagate across the lattice of cells and represent light waves."

Leonard Susskind Black Hole Wars

3.1 Information in interacting systems

The time evolution of a dynamical system may be called non conservative and stochastic if it generates 'hidden information' at a nonzero entropy rate at different time scales and spatial resolution. For systems consisting of more than one component (e.g. spin), each individual component will contribute to the generation of information. Interacting systems even exchange information among each other. Especially for systems comprising a large number of interacting components with an order-disorder phase transition, one would expect to see information theoretic quantities to diverge in analogy to the two point correlation function and the free energy.

3.2 Transfer entropy rates of two interacting systems

Suppose that the the process $\boldsymbol{\sigma}_t$ would consist of two seperate sub-processes,

$$(\boldsymbol{\sigma}_t) \mapsto (\boldsymbol{a}_t, \boldsymbol{b}_t). \tag{3.1}$$

In such a situation the meaning of information and the ability to distinguish between the past and future becomes even more interesting. We can now analyse the information generated by each sub-process $\boldsymbol{a}_t$ and $\boldsymbol{b}_t$ seperately and investigate the interactions among each other.
Assume that from the ensemble (2.1) we have extracted all sequences of the sub-process$\{\boldsymbol{a}_t\}$ and one particular sequence of the other sub-process $\boldsymbol{b}_t$. This allows us to redefine the probability distribution of the sequence $\boldsymbol{a}_t$, i.e.

$$\mathcal{P}(\boldsymbol{a}_t) \mapsto \mathcal{P}(\boldsymbol{a}_t|\boldsymbol{b}_t) \tag{3.2}$$

which is the conditional probability[1].
Due to this redefinition the entropy also changes:

$$\mathbb{H}[\boldsymbol{a}_t] \mapsto -\sum_{\{\boldsymbol{a}_t\}} p(\boldsymbol{a}_t|\boldsymbol{b}_t) \log p(\boldsymbol{a}_t|\boldsymbol{b}_t) \tag{3.4}$$

In case we extract all sequences $\{\boldsymbol{b}_t\}$ from the ensemble, we can take the average of (3.4) and arrive at the *conditional entropy*

$$\begin{aligned} \mathbb{H}[\boldsymbol{a}_t|\boldsymbol{b}_t] &\coloneqq -\sum_{\{\boldsymbol{b}_t\}} p(\boldsymbol{b}_t) \left(\sum_{\{\boldsymbol{a}_t\}} p(\boldsymbol{a}_t|\boldsymbol{b}_t) \log p(\boldsymbol{a}_t|\boldsymbol{b}_t) \right) & (3.5) \\ &= -\Big\langle \log p(\boldsymbol{a}_t|\boldsymbol{b}_t) \Big\rangle_{\{\boldsymbol{\sigma}_t\}}, & (3.6) \end{aligned}$$

also a very important quantitiy in information theory. We note that since

$$\begin{aligned} \mathbb{H}[\boldsymbol{a}_t|\boldsymbol{b}_t] &= \mathbb{H}[\boldsymbol{\sigma}_t] - \mathbb{H}[\boldsymbol{b}_t] & (3.7) \\ \mathbb{H}[\boldsymbol{b}_t|\boldsymbol{a}_t] &= \mathbb{H}[\boldsymbol{\sigma}_t] - \mathbb{H}[\boldsymbol{a}_t] & (3.8) \end{aligned}$$

[1] Here we note that

$$p(\boldsymbol{a}_t|\boldsymbol{b}_t) = \frac{p(\boldsymbol{\sigma}_t)}{p(\boldsymbol{b}_t)} \qquad p(\boldsymbol{b}_t|\boldsymbol{a}_t) = \frac{p(\boldsymbol{\sigma}_t)}{p(\boldsymbol{a}_t)} \tag{3.3}$$

are the joint probability $p(\boldsymbol{\sigma}_t) = p(\boldsymbol{a}_t, \boldsymbol{b}_t)$ as well as the marginals $p(\boldsymbol{a}_t) = \sum_{\{\boldsymbol{b}_t\}} p(\boldsymbol{\sigma}_t)$ and $p(\boldsymbol{b}_t) = \sum_{\{\boldsymbol{a}_t\}} p(\boldsymbol{\sigma}_t)$.

the conditional entropy is a non symmetric measure, i.e.

$$\mathbb{H}[\boldsymbol{a}_t|\boldsymbol{b}_t] - \mathbb{H}[\boldsymbol{b}_t|\boldsymbol{a}_t] = \mathbb{H}[\boldsymbol{a}_t] - \mathbb{H}[\boldsymbol{b}_t]. \tag{3.9}$$

The conditional entropy measures the reduction of uncertaintiy about the path $\boldsymbol{a}_t$ or $\boldsymbol{b}_t$ if the additional information of one of the latter is taken into account. In a similar fashion having observed the ensemble already for a longer time can only decrease the uncertainty about the future of the path $\boldsymbol{\sigma}_t$ on average. From $p(\boldsymbol{\sigma}_t) \mapsto p(\sigma_{t+dt}|\boldsymbol{\sigma}_t)$, in the stationary regime the conditional entropy is a decreasing funtion of t, i.e.

$$\mathbb{H}[\sigma_{t+dt}|\boldsymbol{\sigma}_t] \leq \mathbb{H}[\sigma_t|\boldsymbol{\sigma}_{t-dt}] \tag{3.10}$$

Thus its limit is the entropy rate (2.19) [14, 42]

$$\lim_{t\to\infty} \frac{1}{t}\mathbb{H}[\sigma_{t+dt}|\boldsymbol{\sigma}_t] = \lim_{t\to\infty} \frac{1}{t}\mathbb{H}[\boldsymbol{\sigma}_t] = \dot{\mathbb{H}}[\boldsymbol{\sigma}]. \tag{3.11}$$

Taking this into account, let us now consider the limit [49]

$$\begin{aligned}
\lim_{t\to\infty} \frac{1}{t}\mathbb{H}[\boldsymbol{a}_t|\boldsymbol{b}_t] &= \lim_{t\to\infty} \frac{1}{t}\Big(\mathbb{H}[\boldsymbol{\sigma}_t] - \mathbb{H}[\boldsymbol{b}_t]\Big) && (3.12)\\
&= \lim_{t\to\infty} \frac{1}{t}\Big(\mathbb{H}[\sigma_{t+dt}|\boldsymbol{\sigma}_t] - \mathbb{H}[b_{t+dt}|\boldsymbol{b}_t]\Big) && (3.13)\\
&= \lim_{t\to\infty} \frac{1}{t}\Big(\mathbb{H}[b_{t+dt}, a_{t+dt}|\boldsymbol{b}_t, \boldsymbol{a}_t] - \mathbb{H}[\boldsymbol{b}_t|\boldsymbol{b}_t]\Big) && (3.14)\\
&= \lim_{t\to\infty} \frac{1}{t}\Big(\mathbb{H}[b_{t+dt}|\boldsymbol{b}_t, \boldsymbol{a}_t] - \mathbb{H}[b_{t+dt}|\boldsymbol{b}_t]\Big) && (3.15)\\
&= \lim_{t\to\infty} \frac{1}{t}\Big(\mathcal{T}_{\boldsymbol{a}\to\boldsymbol{b}}\Big), && (3.16)
\end{aligned}$$

where $\mathcal{T}_{\boldsymbol{a}\to\boldsymbol{b}}$ is the so called *transfer entropy* introduced by Schreiber [52],

$$\begin{aligned}
\mathcal{T}_{\boldsymbol{a}\to\boldsymbol{b}} &:= \mathbb{H}[b_{t+dt}|\boldsymbol{b}_t, \boldsymbol{a}_t] - \mathbb{H}[b_{t+dt}|\boldsymbol{b}_t] && (3.17)\\
&= \left\langle \log \frac{p(b_{t+dt}|\boldsymbol{b}_t, \boldsymbol{a}_t)}{p(b_{t+dt}|\boldsymbol{b}_t)} \right\rangle_{\{\boldsymbol{\sigma}_t\}}, && (3.18)
\end{aligned}$$

where $\mathbb{H}[b_{t+dt}|\boldsymbol{b}_t]$ accounts for the average number of bits needed to encode one additional state b_{t+dt} of the typtical path if all previous states $\boldsymbol{b}_t$ are known. $\mathbb{H}[b_{t+dt}|\boldsymbol{b}_t, \boldsymbol{a}_t]$ is the entropy rate capturing the average number of bits required to represent the system state if additionally the path history $\boldsymbol{a}_t$ is included. Thus if $\boldsymbol{\sigma}_t$ would be a

Markovian process, then $\mathcal{T}_{\boldsymbol{a}\to\boldsymbol{b}} = 0$ only if the coarse grained processes $\boldsymbol{b}$ and $\boldsymbol{a}$ are Markovian as well, i.e.

$$p(b_{t+dt}|\boldsymbol{b}_t, \boldsymbol{a}_t) = p(b_{t+dt}|\boldsymbol{b}_t), \tag{3.19}$$

where the past of $\boldsymbol{a}_t$ has no influence on the transition probabilities of $\boldsymbol{b}_t$. Generally $\mathcal{T}_{\boldsymbol{a}\to\boldsymbol{b}}$ is nonsymmetric since it measures the degree of dependence of the history of $\boldsymbol{a}_t$ on the future path of $\boldsymbol{b}_t$ and not vice versa, i.e.

$$\mathcal{T}_{\boldsymbol{a}\to\boldsymbol{b}} \neq \mathcal{T}_{\boldsymbol{b}\to\boldsymbol{a}}. \tag{3.20}$$

3.3 Mutual information rates of two interacting systems

Finally we want to quantify how predictable the time evolution of the sub-system remains, if we consider both sub-processes $\boldsymbol{a}_t$ and $\boldsymbol{b}_t$ as being absolutely independent processes. Evidently we could construct binary codes for the paths $\boldsymbol{\sigma}_t$, $\boldsymbol{a}_t$ and $\boldsymbol{b}_t$ and compare the average code length. From that we can deduce how much entropy each process generates. If both sub-processes would be independet, we would expect that $\mathbb{H}[\boldsymbol{a}_t] + \mathbb{H}[\boldsymbol{b}_t] = \mathbb{H}[\boldsymbol{\sigma}_t]$. Thus any interaction amongst the sub-processes can be captured by the quantitiy

$$\mathbb{I}[\boldsymbol{b}_t : \boldsymbol{a}_t] \;\coloneqq\; \mathbb{H}[\boldsymbol{a}_t] + \mathbb{H}[\boldsymbol{b}_t] - \mathbb{H}[\boldsymbol{\sigma}_t]. \tag{3.21}$$

Equation (3.21) is called the *mutual information* of the two processes $\boldsymbol{b}_t$ and $\boldsymbol{a}_t$ with joint probability $p(\boldsymbol{\sigma}_t)$. It can be seen as the excess amount of code produced by erroneously assuming that the two sub-systems are independent, i.e., using

$$q(\boldsymbol{\sigma}_t) \coloneqq p(\boldsymbol{b}_t)p(\boldsymbol{a}_t) \tag{3.22}$$

instead of $p(\boldsymbol{\sigma}_t)$. Evidently equation (3.21) is simply the Kullback divergence, i.e.

$$\begin{aligned}
\mathbb{I}[\boldsymbol{b}_t : \boldsymbol{a}_t] &= \mathbb{D}_{\mathrm{KLD}}[\mathcal{P}(\boldsymbol{\sigma}_t) \parallel \mathcal{Q}(\boldsymbol{\sigma}_t)] && (3.23)\\
&= \left\langle \log \frac{p(\boldsymbol{\sigma}_t)}{q(\boldsymbol{\sigma}_t)} \right\rangle_{\{\boldsymbol{\sigma}_t\}} && (3.24)\\
&= \left\langle \log \frac{p(\boldsymbol{\sigma}_t)}{p(\boldsymbol{b}_t)p(\boldsymbol{a}_t)} \right\rangle_{\{\boldsymbol{\sigma}_t\}} \geq 0. && (3.25)
\end{aligned}$$

The mutual information is a suitable measure to quantify the deviation from independence of two correlated processes. Note that $\mathcal{I}[\boldsymbol{b}_t : \boldsymbol{a}_t] = \mathcal{I}[\boldsymbol{a}_t : \boldsymbol{b}_t]$ is symmetric and therefore does not contain any directional sense as compared to the transfer entropy. Similarly to section 2.3 we can define a time dependent rate function of the mutual information, i.e.

$$\dot{\mathbb{I}}[\boldsymbol{b} : \boldsymbol{a}]_t \coloneqq \frac{\mathbb{I}[\boldsymbol{b}_{t+dt} : \boldsymbol{a}_{t+dt}] - \mathbb{I}[\boldsymbol{b}_t : \boldsymbol{a}_t]}{\mathrm{d}t}. \tag{3.26}$$

In order to analyze the mutual information transfer betwenn both sub-systems we can also look at the temporal change of the mutual information if we only aquire new additional information about one process, i.e.

$$\dot{\mathbb{I}}^a[\boldsymbol{b} : \boldsymbol{a}]_t \coloneqq \frac{\mathbb{I}[\boldsymbol{b}_t : \boldsymbol{a}_{t+dt}] - \mathbb{I}[\boldsymbol{b}_t : \boldsymbol{a}_t]}{\mathrm{d}t} \tag{3.27}$$

$$\dot{\mathbb{I}}^b[\boldsymbol{b} : \boldsymbol{a}]_t \coloneqq \frac{\mathbb{I}[\boldsymbol{b}_{t+dt} : \boldsymbol{a}_t] - \mathbb{I}[\boldsymbol{b}_t : \boldsymbol{a}_t]}{\mathrm{d}t} \tag{3.28}$$

These quantities are commonly referred to as the *information flow* or *learning rate* [5, 27] since they quantify how much the dynamics of one sub-system reduced the uncertainty of the others due to their coupling.

Chapter 4

Markovian dynamics

> "What's so bad about losing a bit of information inside a black hole? Then it dawned on us. Losing information is the same as generating entropy. And generating entropy means generating heat. The virtual black holes that Stephen had so blithely postulated would create heat in empty space."
>
> *Leonard Susskind* Black Hole Wars

4.1 Probability distributions and their time evolution laws

Above we have introduced the method of coarse graining, in order to simplify the treatment of complicated, practicably unsolvable deterministic systems. For some simple systems, such approximations can be rigorously derived from the underlying deterministic or quantum dynamics when a mapping is introduced [23]. On the level of of a coarse grainded description it is generally not possible to exactly predict the future sequence of states. It is only possible to predict the probabilities $p(\sigma, t)$ of states of an ensemble of stochastic processes $\{\boldsymbol{\sigma}_t\}$. The probability $p(\sigma, t)$ is the frequency of how often on average an outcome σ will occcur in time in the ensemble $\{\boldsymbol{\sigma}_t\}$. This frequency is of course governed by dynamical laws producing the probabilities of paths $\boldsymbol{\sigma}_t$.

4.2 Continuous time Markovian stochastic processes

For a continous time evolution we consider the so called Master Equation (ME) which is a balance equation of transitions $\sigma \to \sigma'$ and $\sigma' \to \sigma$ between states per unit time, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}p(\sigma,t) = \sum_{\sigma'} W_{\sigma\sigma'}p(\sigma',t) \tag{4.1}$$

or similarly in vector notation

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}_t = \boldsymbol{W}\boldsymbol{p}_t, \tag{4.2}$$

where $\boldsymbol{p}_t \coloneqq (p(\sigma_0,t), p(\sigma_1,t), ..., p(\sigma_n,t))^\intercal$ is a vector representing the probabilities of all states of the system. In equation (4.2) the transition rates $W_{\sigma\sigma'}$ are understood as matrix elements of a so called rate matrix $\boldsymbol{W}$, where for probability conservation it is required that $W_{\sigma\sigma} = -\sum_{\sigma'\neq\sigma} W_{\sigma'\sigma}$.
The solution $\boldsymbol{p}_t$ of the ME describes the stochastic process when the ensemble $\{\boldsymbol{\sigma}_t\}$ evolves in time. As stated before, it is neither possible to predict the actual sequence of transitions nor the times at which jumps happen. They are both random variables characterising the stochastic process $\{\boldsymbol{\sigma}_t\}$. The only quantitiy that can be predicted in a stochastic process are probabilities, e.g. $p(\sigma,t)$ to find the system in state σ at time t.

4.2.1 Master equation on a graph

The equation for the time evolution of a system dictates the neighboring structure of prior and posteriour states in time. The jump process introduced here can be seen as a random walk on a graph $G_{\boldsymbol{W}}$, with each state σ corresponding to one vertex $\sigma \mapsto v \in \boldsymbol{v}$ of a graph. Similar to σ, all vertices are simply indices, i.e. $v \in \mathbb{Z}$. Whenever the system can jump from one state to another, i.e. if $W_{\sigma\sigma'} \neq 0$,then v and v' are connected by an edge $e \in \boldsymbol{e}$. The set of edges $\boldsymbol{e}$ and vertices $\boldsymbol{v}$ build the graph $G_{\boldsymbol{W}}$ assoziated to the transition matrix $\boldsymbol{W}$, i.e.

$$G_{\boldsymbol{W}} = (\boldsymbol{v}, \boldsymbol{e}). \tag{4.3}$$

Schnakenberg has shown that many of the fundamental properties of a nonequilibrium random process can be investigated and understood by carrying out the analysis of the graph associated with the master equation [51].

4.2.2 Detailed balance along edge

While the time evolution of the state of the system is a jump process over the network, the occupation frequency and therefore the probability for the system to be in a state σ can be interpreted as probability currents flowing along the edges of the network. In that sense we define the probability current along an edge e as

$$j_e[\boldsymbol{\sigma}]_t \coloneqq W_{\sigma\sigma'}p(\sigma',t) - W_{\sigma'\sigma}p(\sigma,t) \tag{4.4}$$

If the ME describes an equilibrium process then, as the process $\boldsymbol{\sigma}$ reaches the stationary conditions, the probability currents vanish, i.e.

$$\lim_{t\to\infty} j_e^{eq}[\boldsymbol{\sigma}]_t = 0. \qquad \forall e \tag{4.5}$$

An equivalent condition, known as the *Kolmogorov criterion* [37] (see book by T. Tomé, page 177 [61]), can be expressed even without the knowledge of the steady state. It relies on the notion of closed cyclic sequence $c \coloneqq (\sigma_1, \sigma_2, ...\sigma_n, \sigma_1)^\intercal$ and their reverse c^R. With the definition of $W(c) \coloneqq W_{\sigma_1\sigma_n}W_{\sigma_{n-1}\sigma_{n-2}}...W_{\sigma_2\sigma_1}$ the Kolmogorov criterion, for a Markovian stochastic process to be an equilibrium procces, reads:

$$\forall c \in \boldsymbol{c}: \qquad a_c[\boldsymbol{\sigma}] \coloneqq \log\frac{W(c)}{W(c^R)} = 0, \tag{4.6}$$

where $\boldsymbol{c}$ holds all possible cycles, that can be constructed by the edge set $\boldsymbol{e}$. $a_c[\boldsymbol{\sigma}]$ is often referred to as the *cycle-affinity*. Equation (4.6) is a necessary condition for statistical time reversal symmetry to hold. If this condition is violated for any cycle, i.e.

$$\exists c \in \boldsymbol{c}: \qquad a_c[\boldsymbol{\sigma}] \neq 0, \tag{4.7}$$

the system is in a NESS characterized by the presence of non vanishing probability currents (4.4) and non zero entropy production rate as we will se below.
Utilizing the condition (4.10) we can introduce a potential function Ω_σ via

$$p^{eq}(\sigma) = \frac{1}{Z}\exp(\Omega_\sigma), \qquad Z \coloneqq \sum_{\boldsymbol{\sigma}}\exp(\Omega_\sigma). \tag{4.8}$$

We can identify the potential Ω_σ with an effective Hamiltonian and Z with the associated canonical partition function. In this fashion, a stochastic process breaking the detailed balance condition describes a physical system relaxing towards thermal equilibrium (4.8). The detailed balance condition can be expressed as

$$g_e^{eq}[\boldsymbol{\sigma}] \coloneqq \frac{W_{\sigma\sigma'}}{W_{\sigma'\sigma}} = \frac{p^{eq}(\sigma)}{p^{eq}(\sigma')} = \exp(\Omega_\sigma - \Omega_{\sigma'}). \tag{4.9}$$

4.2.3 Entropy production and entropy flow along edges

In this section we again start from the perspective that we want to quantify whether and how much a system, described by a continuous time ME, breaks the statistical time reversal symmetry. Above we have seen, that this can be investigate by looking at the probability of typical paths and their reverse of the stochastic process. Statistical time reversal symmetry of a Markovian process described by a ME is broken if $W_{\sigma\sigma'}p_{\sigma'}(t) \neq W_{\sigma'\sigma}p_{\sigma}(t)$ along any edge, precisely [60]

$$j_e[\boldsymbol{\sigma}]_t \neq 0. \tag{4.10}$$

In that sense and in analogy to above (c.f. equation (2.9)), the breaking of statistical time reversal symmetry can be quantified by

$$\log j_e[\boldsymbol{\sigma}]_t = \log \frac{W_{\sigma\sigma'}p_{\sigma'}(t)}{W_{\sigma'\sigma}p_{\sigma}(t)}, \tag{4.11}$$

a logarithmic measure the breaking of detailed balance, often called *affinity*. We note here that the product of (4.4) and (4.11) is always a non negative quantitiy, i.e.

$$j_e[\boldsymbol{\sigma}]_t \log j_e[\boldsymbol{\sigma}]_t = \quad \geq \quad 0. \qquad \forall e \wedge \forall t \tag{4.12}$$

We can split the time derivative of Shannons entropy as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{H}[\boldsymbol{\sigma}]_t \quad = \quad \dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}]_t + \dot{\mathcal{S}}_{env}[\boldsymbol{\sigma}]_t \tag{4.13}$$

where (see M. Esposito [64] and C. van Boerek [63])

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}]_t \quad \coloneqq \quad \sum_{\boldsymbol{e}} j_e[\boldsymbol{\sigma}]_t \log j_e[\boldsymbol{\sigma}]_t \geq 0 \tag{4.14}$$

is the *entropy production rate* of the system and

$$\dot{\mathcal{S}}_{env}[\boldsymbol{\sigma}]_t \quad \coloneqq \quad -\sum_{\boldsymbol{e}} j_e[\boldsymbol{\sigma}]_t \log g_e[\boldsymbol{\sigma}]^{eq} \tag{4.15}$$

the *entropy flow* between the environment and the system. The entropy production rate (4.14), as required, is always non-negative and vanishes at equilibirum. If the ME describes an equilibirum process, then the relaxation of an initial distribution $\boldsymbol{p}_0$ represents an irreversible process which terminates in the limit $t \to \infty$, where detailed balance holds. In this case the entropy production rate vanishes, i.e. $\dot{\mathcal{S}}_i[\boldsymbol{\sigma}]_\infty = 0$,

but surely $\int_0^\infty \mathrm{d}t\dot{\mathcal{S}}_i[\boldsymbol{\sigma}]_t \geq 0$. Consequently the entropy production is a measure of the deviation of a stochastic process from a equilibirum distribution either as a NESS property or restricted to a finite relaxation time interval. Generally, stochastic processes which do not satisfy detailed balance have a non vanishing entropy production. Though everything looks straightforward, computing the entropy production rate via Eq. (4.14) is often unfeasible for complicated system. First of all the computation requires that all possible states of $\boldsymbol{\sigma}$ can be identified, that all rates are available, and that the steady-state can be computed. Even the simplest of these tasks, identifying the sets of states, is already a demanding task for a simple 2D-Ising model where the state space is made up by all different configurations the spin-system can be in.

4.2.4 Relaxation properties

A given initial probability of outcomes $\boldsymbol{p}_0 \coloneqq (p(\sigma_0, 0), ..., p(\sigma_n, 0))^\intercal$ evolves as a solution to the linear differential equation (4.2),

$$\boldsymbol{p}_t = e^{\boldsymbol{W}t}\boldsymbol{p}_0. \tag{4.16}$$

This reflects that (4.2) is a set of coupled first-order linear differential equations with constant coefficients. Equation (4.16) can be written in a time series expansion, excplicitly

$$\boldsymbol{p}_t = (\mathbf{1} + \boldsymbol{W}t + \frac{t^2}{2!}\boldsymbol{W}^2 + \frac{t^3}{3!}\boldsymbol{W}^3 + ...)\boldsymbol{p}_0. \tag{4.17}$$

If $\boldsymbol{W}^s$ is determined for every power $s \in \mathbb{N}$, then (4.17) is a solution to the master equation (4.2). The solution of the master equation can also be obtained from the eigenvectors and eigenvalues of the rate matrix $\boldsymbol{W}$ [61]. We denote the left and right eigenvectors by $\boldsymbol{l}_k$ and $\boldsymbol{r}_k$ and the corresponding eigenvalues by λ_k, i.e.

$$\boldsymbol{W}\boldsymbol{r}_k = \lambda_k \qquad \boldsymbol{l}_k\boldsymbol{W} = \lambda_k. \tag{4.18}$$

We assume that the eigenvectors form a complete and orthogonal set. Therefore we can consider the following expansion

$$e^{\boldsymbol{W}t} = e^{\boldsymbol{W}t}\mathbf{1} = e^{\boldsymbol{W}t}\sum_k \boldsymbol{r}_k\boldsymbol{l}_k = \sum_k e^{\lambda_k t}\boldsymbol{r}_k\boldsymbol{l}_k. \tag{4.19}$$

With the use of the steady state solution $\boldsymbol{W}\overline{\boldsymbol{p}} = 0$ we can rewirte (4.16) as

$$\boldsymbol{p}_t = \overline{\boldsymbol{p}} + \sum_{k\neq 0} e^{\lambda_k t}\boldsymbol{r}_k\boldsymbol{l}_k\boldsymbol{p}_0, \tag{4.20}$$

since $\boldsymbol{W}$ fullfills the requirements of the *Perron-Frobenius* theorem [61]. Therefore the eigenvalues of the matrix $\boldsymbol{W}$ are negative or zero and it follows that $\lim_{t\to\infty} \boldsymbol{p}_t = \bar{\boldsymbol{p}}$. Since there are usually large gaps in the eigenvalue spectrum we can obtain a good approximation to the long-term behavior by keeping only terms with the smallest eigenvalues $|\lambda_k|$.

For MEs and in fact any stochastic process generated by a linear map, the time derivative of the relative entropy of any two initial distributions $\boldsymbol{p}_0$ and $\boldsymbol{p}'_0$ [63, 55], i.e.

$$\frac{\mathrm{d}\mathbb{D}_{\mathrm{KLD}}[\boldsymbol{p}_0 \parallel \boldsymbol{p}'_0]}{\mathrm{d}t} \leq 0, \tag{4.21}$$

is a decreasing function in time. Thus as time evolves, two initially distingusishable distributions become indistinguishable. Especially [51]

$$\frac{\mathrm{d}\mathbb{D}_{\mathrm{KLD}}[\boldsymbol{p}_0 \parallel \bar{\boldsymbol{p}}]}{\mathrm{d}t} \leq \lim_{t\to\infty} \frac{\mathrm{d}\mathbb{D}_{\mathrm{KLD}}[\boldsymbol{p}_0 \parallel \bar{\boldsymbol{p}}]}{\mathrm{d}t} = 0. \tag{4.22}$$

In that sense information is lost during the relaxation process towards the stationary state.

4.2.5 Diagrammatic construction of the steady state

In order to obtain a steady state solution of (4.2), one would have to solve the linear equation

$$\boldsymbol{W}\bar{\boldsymbol{p}} = 0. \tag{4.23}$$

under the constraint that $\bar{\boldsymbol{p}}$ is a probability distribution. The existence of such a solution is guaranteed, since it can explicitly be constructed by graph theoretical methods [51]. The representation of a ME by its basic graph (4.3) allows for an elegant and straight forward method to construct the steady state vector $\bar{\boldsymbol{p}}$ [51]. A detailed mathematical description of this method can be found in the thesis of M. Polettini [47]. Formally, the method is based on the definition of a *maximal tree* $T(G)$, which is a subgraph of the graph G satisfying the following properties:

- $T(G)$ contains all the vertices of G;
- $T(G)$ is connected;

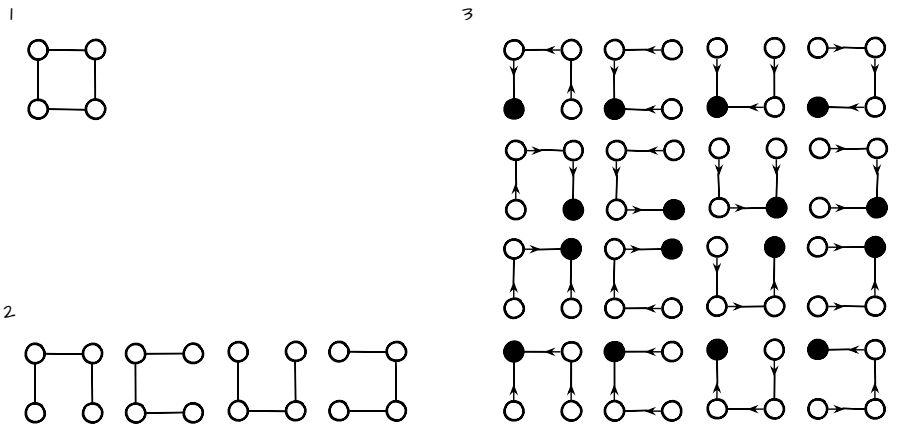

Figure 4.1: (1) Example of a graph G. (2) G has 4 maximal trees. (3) G has 16 i-oriented (i marked blacck) maximal trees

- $T(G)$ contains no circuit, i.e. no cyclic sequence of edges.

For a graph G there are many ways to construct a maximal tree $T(G)$. We label each element of the complete set of maximal trees by $T^{(\mu)}(G)$, with $\mu = 1, 2, ...M$. The total number M of maximal trees closely depends on the topological structure of G and on its number of edges. For each of the maximal trees $T^{(\mu)}(G)$ we construct all its i-directed version $T_i^{(\mu)}(G)$ by orienting all edges of $T^{(\mu)}(G)$ along the path to the vertex with label i (see figure 4.1).
With the evaluation function

$$\pi_i\left[\boldsymbol{\sigma}\right] \coloneqq \sum_{\mu} \left(\prod_{e \in T_i^{(\mu)}(G)} W_{\sigma\sigma'} \right) \tag{4.24}$$

which multiplies all transition rates for jumps along the orientation of the edges of the i-oriented maximal tree, and sums over all possible maximal trees μ. Finally the steady state solution of the ME (4.2)is given as

$$\overline{p}(\sigma) = \frac{\pi_\sigma\left[\boldsymbol{\sigma}\right]}{\sum_\sigma \pi_z\left[\boldsymbol{\sigma}\right]} \tag{4.25}$$

4.2.6 Independent probability currents at steady state

As introduced in section 4.2.2, equilibrium conditions can be rephrased in terms of the cyclic paths of the stochastic process with a vanishing cycle-affinity (4.6). J.

Schnakenberg found that non equilibrium constraints in turn are related to non-zero affinities of special subsets of the cyclic paths [51]. A graph G usually has a large number of cyclic paths $\boldsymbol{c}$. However, not all cyclic paths are independent. They can be expressed by a linear combination of a smaller subset of cycles, called the *fundamental cycle set* $\boldsymbol{c_l}$, which plays the role of a basis in the space of cycles. Here we introduce the basic methodology to find such a set $\boldsymbol{c_l}$.
All edges of the graph G associated with the transition matrix $\boldsymbol{W}$ are assigned with an arbitrary reference orientation, here

$$o_e := \begin{cases} +1 & \sigma < \sigma' \\ -1 & \sigma > \sigma' \end{cases} \tag{4.26}$$

All edges $\boldsymbol{l} \subset \boldsymbol{e}$ not belonging to a maximal tree $T(G)$ (see figure 4.2 (2)) are called *chords*. Adding such a chord l to the maximal tree results in a subgraph that contains exactly one cycle (see figure 4.2 (3)). A circular orientation

$$o_l := +1 \tag{4.27}$$

is assigned to the edges of the fundamental cycles associated to $T(G)$, defining the fundamental set of circuits $\boldsymbol{c_l}$. Once the fundamental set is found, we finally need one more tool to check whether the orientation of edges in the graph G is parallel or antiparallel to any constructed subgraph f of G. Therefore we define

$$\delta_e(f) := \begin{cases} +1 & \text{if } e \text{ and } f \text{ are parallel, i.e. } o_e o_f = +1 \\ -1 & \text{if } e \text{ and } f \text{ are antiparallel, i.e. } o_e o_f = -1 \\ 0 & \text{if } e \text{ is not in } f \end{cases} \tag{4.28}$$

Finally the mean current traversing the edge e can be expressed as [51, 23]

$$j_e[\boldsymbol{\sigma}] = \sum_l \delta_e(c_l) j_l[\boldsymbol{\sigma}], \tag{4.29}$$

which is the very important result that the probability currents of the system are dependent and their magnitude is confined to the fundamental set of cycles and given by the probability current of the chord. This is in close analogy to *Kirchhoffs current law* and the probability conservation of the ME.

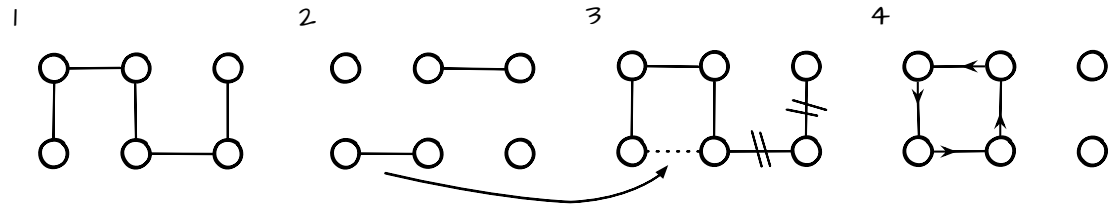

Figure 4.2: (1) Example of a maximal tree $T(G)$. (2) The chords l associated to $T(G)$. (3) Adding one of the chords to the maximal tree creates a closed cycle. Edges not belonging to this cycle are erased (4) The cycle is assigned a cyclic orientation.

4.2.7 Steady state entropy production of fundamental cycles

The above machinery enables us to define the affinities and of an oriented cycle [51, 23], i.e.

$$\begin{aligned} a_c[\boldsymbol{\sigma}] &\coloneqq \sum_e \delta_e(c) \log j_e[\boldsymbol{\sigma}] && (4.30)\\ &= \log \prod_e \delta_e(c) \frac{W_{\sigma\sigma'}\overline{p}(\sigma')}{W_{\sigma'\sigma}\overline{p}(\sigma)} && (4.31)\\ &= \log \prod_e \delta_e(c) \frac{W_{\sigma\sigma'}}{W_{\sigma'\sigma}} && (4.32)\\ &= \log \prod_e \delta_e(c) g_e[\boldsymbol{\sigma}]^{eq} && (4.33)\\ &= \log \frac{W(c)}{W(c^R)}, && (4.34) \end{aligned}$$

since indeed the terms $\log \frac{\overline{p}_{\sigma'}}{\overline{p}_\sigma}$ cancel each other along a closed path. Thanks to the fundamental cycles, each cycle can now be decomposed. The affinity (4.34) of an arbitrary cycle c can be expressed as a combination of the fundamental cycles

$$a_c[\boldsymbol{\sigma}] = \sum_l \delta_l(c) a_l[\boldsymbol{\sigma}], \qquad \text{with} \qquad a_l[\boldsymbol{\sigma}] \coloneqq \log \frac{W(c_l)}{W(c_l^R)}. \tag{4.35}$$

We can now finally express the entropy production rate, at steady state, in terms of independent currents and affinities along the fundamental cycles sets

$$\begin{aligned}
\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}] &= \sum_e j_e[\boldsymbol{\sigma}] \log j_e[\boldsymbol{\sigma}] && (4.36)\\
&\underset{(4.13)}{=} \sum_e j_e[\boldsymbol{\sigma}] \log g_e[\boldsymbol{\sigma}]^{eq} && (4.37)\\
&\underset{(4.29)}{=} \sum_l j_l[\boldsymbol{\sigma}] \sum_e \delta_e(c_l) \log g_e[\boldsymbol{\sigma}]^{eq} && (4.38)\\
&\underset{(4.34)}{=} \sum_l j_l[\boldsymbol{\sigma}] \log \frac{W(c_l)}{W(c_l^R)} && (4.39)\\
&\underset{(4.35)}{=} \sum_l j_l[\boldsymbol{\sigma}] a_l[\boldsymbol{\sigma}] \geq 0 && (4.40)
\end{aligned}$$

4.2.8 Dynamical randomness and entropy production

Even though the definition of (4.14) as the entropy production rate only generates none-negative values, we need to show a closer connection to the breaking of statistical time reversal symmetry, as introduced in the chapters above. It can be shown that

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}] = \dot{\mathbb{H}}^R[\boldsymbol{\sigma}] - \dot{\mathbb{H}}[\boldsymbol{\sigma}], \tag{4.41}$$

holds. The probability that, during a time interval dt, the system undergoes a transition from state σ' to state σ is, according to equation (4.16), given by

$$p_{dt}(\sigma|\sigma') \coloneqq \left[e^{\boldsymbol{W}dt}\right]_{\sigma\sigma'}. \tag{4.42}$$

For very short times dt, keeping terms up to the first order is sufficient, i.e.

$$p_{dt}(\sigma|\sigma') = \delta_{\sigma\sigma'} + W_{\sigma\sigma'}dt + \mathcal{O}(dt^2). \tag{4.43}$$

With the results shown for discrete time stochasic processes (see Appendix (C.2.2)) we have the result that

$$\begin{aligned}
\dot{\mathbb{H}}^R[\boldsymbol{\sigma}] - \dot{\mathbb{H}}[\boldsymbol{\sigma}] &= -\sum_{\sigma\sigma'} \overline{p}(\sigma)\Big(p_{dt}(\sigma'|\sigma)\log p_{dt}(\sigma|\sigma')\Big) + \sum_{\sigma\sigma'} \overline{p}(\sigma)\Big(p_{dt}(\sigma'|\sigma)\log p_{dt}(\sigma'|\sigma)\Big)\\
&= \frac{1}{2}\sum_{\sigma\sigma'} p_{dt}(\sigma'|\sigma)\overline{p}(\sigma) - p_{dt}(\sigma|\sigma')\overline{p}(\sigma')\log\frac{p_{dt}(\sigma'|\sigma)\overline{p}(\sigma)}{p_{dt}(\sigma|\sigma')\overline{p}(\sigma')} \qquad (4.44)
\end{aligned}$$

$$= \frac{1}{2}\sum_{\sigma\sigma'} W_{\sigma'\sigma}\overline{p}(\sigma) - W_{\sigma\sigma'}\overline{p}(\sigma')\log\frac{W_{\sigma'\sigma}\overline{p}(\sigma)}{W_{\sigma'\sigma}\overline{p}(\sigma')} + \mathcal{O}(dt) \tag{4.45}$$

$$= \sum_{\boldsymbol{e}} j_e[\boldsymbol{\sigma}]\log j_e[\boldsymbol{\sigma}] + \mathcal{O}(dt) \tag{4.46}$$

$$= \lim_{t\to\infty}\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}]_t \tag{4.47}$$

which proofs equation (4.41) [22]. Lets refer back to equation (2.16), namely

$$\Big\langle s(\boldsymbol{\sigma}_t)\Big\rangle_{\{\boldsymbol{\sigma}_t\}} = \mathbb{D}_{\mathrm{KLD}}[\mathcal{P}(\boldsymbol{\sigma}_t) \parallel \mathcal{P}(\boldsymbol{\sigma}_t^R)] \geq 0, \tag{4.48}$$

the information theoretic measure of statistical temporal asymmetry, which represents the expected number of extra bits that are needed to encode the state of the ensemble $\{\boldsymbol{\sigma}_t\}$ if we use a code based on $\{\boldsymbol{\sigma}_t^R\}$. With equation (2.23) we now immediately see an intimate relationship between the information theoretic entropic rate, introduced above and the entropy prodction of a stationary Markov process, i.e.

$$\lim_{t\to\infty}\frac{1}{t}\Big\langle s(\boldsymbol{\sigma}_t)\Big\rangle_{\{\boldsymbol{\sigma}_t\}} = \sum_{\boldsymbol{e}} j_e[\boldsymbol{\sigma}]\log j_e[\boldsymbol{\sigma}] + \mathcal{O}(dt). \tag{4.49}$$

In a thermodynamic context this information theoretic measure is interpreted as the dissipation rate of the system which is observed.

4.2.9 Stochatic Thermodynamics

In the canonical ensemble, the physical system described by a Hamiltonian $\mathcal{H}(\boldsymbol{\sigma})$ is in contact with a single heat bath, which provides a constant temperature β. The corresponding steady state is accordingly given by the Boltzmann faktors

$$p^{eq}(\sigma) = \frac{1}{Z(\beta)}\exp\left(-\beta\mathcal{H}(\sigma)\right) \tag{4.50}$$

Transition rates satisfying detailed balance therfore obey

$$\log\frac{W_{\sigma\sigma'}}{W_{\sigma'\sigma}} = \beta\mathcal{H}(\sigma') - \beta\mathcal{H}(\sigma) \tag{4.51}$$

which is often also called the condition of *local detailed balance*. When transitions between system states are mediated by thermodynamic reservoirs, the statistical time irreversibility of the process is a measure of the physical dissipation. For a system coupled to several reservoirs at different temperature detailed balance generally doesn't

hold. Such a setup is typically modelled by a ME accounting for all transitions which can be mediated by these reservoirs as

$$\frac{d}{dt}\boldsymbol{p}_t = \sum_{\nu} \boldsymbol{W}^{(\nu)}\boldsymbol{p}_t, \tag{4.52}$$

where ν labels the different reservoirs for which the transition matrix satisfies local detailed balance. The steady state will be characterized by the presence of non vanishing probability currents and a positive entropy production.
Since the Markov dynamics do not specifically model the exact state of the thermodynamic reservoirs, i.e. how much energy and how many particles there are in the baths, the impact of the reservoirs appears implicitly in the rate matrices $\boldsymbol{W}^{(\nu)}$. While the rates satisfy local detailed balance with the reservoirs this does not imply that the Markov dynamics obey detailed balance. When coupled to multiple reservoirs with incompatible equilibrium states, the Markov dynamics break detailed balance, i.e.

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}]_t = \sum_{\boldsymbol{e},\nu} j_e^{(\nu)}[\boldsymbol{\sigma}]_t \log j_e^{(\nu)}[\boldsymbol{\sigma}]_t > 0 \tag{4.53}$$

and the entropy flow relates the rates to the flow of entropy (heat, particles etc.) from the reservoirs

$$\begin{aligned} \dot{\mathcal{S}}_e[\boldsymbol{\sigma}]_t &= -\sum_{\boldsymbol{e},\nu} j_e^{(\nu)}[\boldsymbol{\sigma}]_t \log g_e^{(\nu)}[\boldsymbol{\sigma}]^{eq} && (4.54)\\ &= \sum_{\boldsymbol{e},\nu} j_e^{(\nu)}[\boldsymbol{\sigma}]_t \Big(\Omega_{\sigma'}^{(\nu)} - \Omega_{\sigma}^{(\nu)}\Big) \neq 0. && (4.55) \end{aligned}$$

A thermodynamic study of the single molecular enzyme trypthophan synthase will be presented below in chapter 6. This study goes into details of the thermodynamic interpretation of Markov processes, where the rates obey local detailed balance and many reservoirs are involved. A good introductory overview to Stochastic Thermodynamics is also given by M. Esposito [64] and C. van Boerek [63].

Part II

Information Thermodynamics

Chapter 5

Interacting Markovian systems

In section 3.1 we saw that in systems consisting of more than one component (e.g. spin), each individual component will contribute to information production. In fact, interacting systems even exchange information among each other. In the following we analyse a stochastic processes $\boldsymbol{\sigma}_t$ which consist of two seperate sub-processes,

$$(\boldsymbol{\sigma}_t) \mapsto (\boldsymbol{a}_t, \boldsymbol{b}_t). \tag{5.1}$$

In particular, it is first assumed that the fluctuations in each subsystem are independent of each other. A system satisfying this condition is generally called a bipartite system. For such a system we can quantify the information generated and exchanged by each sub-process $\boldsymbol{a}_t$ and $\boldsymbol{b}_t$ seperately.

5.1 Information flow between subsystems with bipartite coupling

A formalism introduced by Horowitz and Esposito [30] enables us to characterize how information flow influences the probability currents of a system that evolves according to a bipartite Markovian master equation,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}p(a,b,t) &= \sum_{a'b'} W^{bb'}_{aa'} p(a',b',t) && (5.2)\\
&= \sum_{a'} W^{b}_{aa'} p(a',b,t) + \sum_{b'} W^{bb'}_{a} p(a,b',t) && (5.3)\\
&= \sum_{a'\neq a} j^{b}_{aa'}(t) + \sum_{b'\neq b} j^{bb'}_{a}(t). && (5.4)
\end{aligned}$$

Thus the probability currents of the joint system can be split into seperate currents of the subsystems. Transition rates $W^b_{aa'}$ within the subsystem $\boldsymbol{a}$ are affected by the state of the subsystem $\boldsymbol{b}$ as indicated by the index b and vice versa for subsystem $\boldsymbol{b}$. We call transitions affected by the other subsysten *feedback-transitions*. Equivalently to the ME (5.4) we can split the entropy production of the joint system as

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}]_t = \underbrace{\sum_{a'\geq a;b} j^b_{aa'}(t)\log\frac{j^b_{aa'}(t)}{j^b_{a'a}(t)}}_{\coloneqq\Sigma^{\boldsymbol{a}}(t)} + \underbrace{\sum_{b'\geq b;a} j^{bb'}_a(t)\log\frac{j^{bb'}_a(t)}{j^{b'b}_a(t)}}_{\coloneqq\Sigma^{\boldsymbol{b}}(t)} \tag{5.5}$$

where $\Sigma^{\boldsymbol{a}}(t)$ and $\Sigma^{\boldsymbol{b}}(t)$ are the contributions to the entropy production based on the information from the joint probability distribution. We define $\sigma^{\boldsymbol{a}}(t)$ and $\sigma^{\boldsymbol{b}}(t)$ as the entropy production based on the marginal distributions, i.e.

$$\sigma^{\boldsymbol{a}}(t) = \sum_{a'\geq a} j_{aa'}(t)\log\frac{j_{aa'}(t)}{j_{a'a}(t)} \tag{5.6}$$

$$\sigma^{\boldsymbol{b}}(t) = \sum_{b'\geq b} j_{bb'}(t)\log\frac{j_{bb'}(t)}{j_{b'b}(t)} \tag{5.7}$$

where $j_{aa'}(t) \coloneqq W^b_{aa'}\sum_b p(a',b,t) - W^b_{a'a}\sum_b p(a,b,t)$ and $j_{bb'}(t)$ similarly. In their paper Horowitz and Esposito show that [30]

$$\Sigma^{\boldsymbol{a}}(t) = \sigma^{\boldsymbol{a}}(t) + \dot{\mathbb{I}}^{\boldsymbol{a}}(t) \tag{5.8}$$

$$\Sigma^{\boldsymbol{b}}(t) = \sigma^{\boldsymbol{b}}(t) + \dot{\mathbb{I}}^{\boldsymbol{b}}(t) \tag{5.9}$$

where the time derivative of the mutual information has been split as (c.f. (3.27))

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{I}[b:a]_t = \underbrace{\sum_{a'\geq a;b} j^b_{aa'}(t)\log\frac{p(b|a,t)}{p(b|a',t)}}_{\coloneqq\dot{\mathbb{I}}^{\boldsymbol{a}}(t)} + \underbrace{\sum_{b'\geq b;a} j^{bb'}_a(t)\log\frac{p(a|b,t)}{p(a|b',t)}}_{\coloneqq\dot{\mathbb{I}}^{\boldsymbol{b}}(t)}. \tag{5.10}$$

Evidently in the limit of $t\to\infty$ the information flow of both sub-systems are required to cancel each other, i.e.

$$\lim_{t\to\infty}\dot{\mathbb{I}}^{\boldsymbol{a}}(t) + \dot{\mathbb{I}}^{\boldsymbol{b}}(t) = 0 \tag{5.11}$$

5.2 Partially non-bipartite interaction

We assume that, besides the bipartite feedback transitions, there are also cross-transitions where the states of both subsystems change simultaneously, i.e. $(a,b) \leftrightarrow (a',b')$ with $a \neq a'$ and $b \neq b'$ as indicated by a transition rate $W_{aa'}^{bb'}$. O.b.d.A. we assume $t \to \infty$.

For later convenience we define the mutual information $i(a,b)$ for a pair of states (a,b) as

$$i(a,b) = \ln \frac{p(a,b)}{p(a)p(b)}. \tag{5.12}$$

The average of $i(a,b)$ over all states (a,b) yields the mutual information of the entire system. The time derivate of the mutual information can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{I}[\boldsymbol{b}:\boldsymbol{a}]_t = \sum_{a' \geq a} \iota_{a,a'}^{A} + \sum_{b' \geq b} \iota_{b,b'}^{B} + \sum_{a' \geq a, b' \geq b} \iota_{a,a'}^{b,b'} \tag{5.13}$$

where the first two sums are taken over feedback-transitions in subsystems $\boldsymbol{a}$ or $\boldsymbol{b}$ and the last sum includes all cross-transitions in the considered system. We have

$$\begin{aligned} \iota_{a,a'}^{\boldsymbol{a}} &= \sum_b j_{a,a'}^{b} \left[i(a,b) - i(a',b) \right] \\ &= \sum_b j_{a,a'}^{b} \ln \frac{p(b|a)}{p(b|a')}, \end{aligned} \tag{5.14}$$

$$\begin{aligned} \iota_{b,b'}^{\boldsymbol{b}} &= \sum_a j_{a}^{b,b'} \left[i(a,b) - i(a,b') \right] \\ &= \sum_a j_{a}^{b,b'} \ln \frac{p(a|b)}{p(a|b')}, \end{aligned} \tag{5.15}$$

$$\iota_{a,a'}^{b,b'} = j_{a,a'}^{b,b'} \left[i(a,b) - i(a',b') \right]. \tag{5.16}$$

Thus $\iota_{a,a'}^{\boldsymbol{a}}$ yields the contribution to the total rate of change of mutual information due to the feedback-transition between the states a and a' that takes place in the subsystem $\boldsymbol{a}$ and is affected by the subsystem $\boldsymbol{b}$. A similar interpretation holds for $\iota_{b,b'}^{B}$. The term $\iota_{a,a'}^{b,b'}$ represents the contribution to the total rate of change of mutual information due to the cross-transition between (a,b) and (a',b'), with $a \neq a'$ and $b \neq b'$, that directly connects the two subsystems $\boldsymbol{a}$ and $\boldsymbol{b}$.

Now we derive the influence of the coupling through regulatory and cross-transitions on each of the entire subsystems $\boldsymbol{a}$ and $\boldsymbol{b}$. Therefore, we consider the amount of

entropy $\Sigma^{\boldsymbol{a}}$ produced per unit time in the transitions that change the state of the $\boldsymbol{a}$ subsystem.

$$\Sigma^{\boldsymbol{a}} = \sum_{a\geq a';b} j^{b}_{a,a'} \log \frac{W^{b}_{a,a'} p(a',b)}{W^{b}_{a',a} p(a,b)} + \sum_{a\geq a',b\geq b'} j^{b,b'}_{a,a'} \log \frac{W^{b,b'}_{a,a'} p(a',b')}{W^{b,b'}_{a,a'} p(a,b)}. \tag{5.17}$$

In a similar way, the amount of entropy $\Sigma^{\boldsymbol{b}}$ produced in the $\boldsymbol{b}$ subsystem can be found

$$\Sigma^{\boldsymbol{b}} = \sum_{b\geq b';a} j^{b,b'}_{a} \log \frac{W^{b,b'}_{a} p(a,b')}{W^{b',b}_{a} p(a,b)} + \sum_{a\geq a',b\geq b'} j^{b,b'}_{a,a'} \log \frac{W^{b,b'}_{a,a'} p(a',b')}{W^{b,b'}_{a,a'} p(a,b)}. \tag{5.18}$$

Suppose now that we observe the subsystem $\boldsymbol{a}$ without the knowledge of the states of the subsystem $\boldsymbol{b}$, i.e. we have no access to the joint probability distribution $p(a,b)$ and use instead of it the probability distribution $p(a)$. Proceeding in this way, the *apparent* entropy production $\sigma^{\boldsymbol{a}}$ assigned to the subsystem $\boldsymbol{a}$ is obtained

$$\sigma^{\boldsymbol{a}} = \sum_{a'\geq a;b} j^{b}_{a,a'} \log \frac{W^{b}_{a,a'} p(a')}{W^{b}_{a',a} p(a)} + \sum_{a\geq a';b\geq b'} j^{b,b'}_{a,a'} \log \frac{W^{b,b'}_{a,a'} p(a')}{W^{b',b}_{a',a} p(a)}. \tag{5.19}$$

Similarly, we have

$$\sigma^{\boldsymbol{b}} = \sum_{b\geq b';a} j^{b,b'}_{a} \log \frac{W^{b,b'}_{a} p(b')}{W^{b',b}_{a} p(b)} + \sum_{a\geq a',b\geq b'} j^{b,b'}_{a,a'} \log \frac{W^{b,b'}_{a,a'} p(b')}{W^{b',b}_{a',a} p(b)}. \tag{5.20}$$

The real entropy production rates $\Sigma^{\boldsymbol{a}}$ and $\Sigma^{\boldsymbol{b}}$ are always non-negative, whereas the apparent entropy production rates $\sigma^{\boldsymbol{a}}$ and $\sigma^{\boldsymbol{b}}$ can also be negative [30]. The influence on the entropy production of system $\boldsymbol{a}$ (respectively, $\boldsymbol{b}$) through coupling to the whole system is then given by the difference between the apparent and total entropy production. In analogy to (5.9) we define

$$\dot{\mathbb{I}}^{\boldsymbol{a}} := \sigma^{A} - \Sigma^{A} \tag{5.21}$$

$$\dot{\mathbb{I}}^{\boldsymbol{b}} := \sigma^{B} - \Sigma^{B}. \tag{5.22}$$

Which can be expressed as

$$\dot{\mathbb{I}}^{\boldsymbol{a}} = \sum_{a\geq a';b} j^{b}_{a,a'} \log \frac{p(b|a)}{p(b|a')} + \sum_{a\geq a';b\geq b'} j^{b,b'}_{a,a'} \log \frac{p(b|a)}{p(b'|a')}, \tag{5.23}$$

$$\dot{\mathbb{I}}^{\boldsymbol{b}} = \sum_{b\geq b';a} j^{b,b'}_{a} \log \frac{p(a|b)}{p(a|b')} + \sum_{a\geq a';b\geq b'} j^{b,b'}_{a,a'} \log \frac{p(a|b)}{p(a'|b')}.$$

Note that $\dot{\mathbb{I}}^{\boldsymbol{a}}$ and $\dot{\mathbb{I}}^{\boldsymbol{b}}$ have contributions from terms $\iota^{A}_{a,a'}$ and $\iota^{B}_{b,b'}$ defined in equations (5.14) and (5.15) and used in the splitting of $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{I}[\boldsymbol{b}:\boldsymbol{a}]$ in equation (5.13). In addition, they also include cross-terms originating from non-bipartite transitions. Using $\dot{\mathbb{I}}^{\boldsymbol{a}}$ and $\dot{\mathbb{I}}^{\boldsymbol{b}}$, equation (5.13) for the rate of change of mutual information can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{I}[\boldsymbol{b}:\boldsymbol{a}] = \dot{\mathbb{I}}^{\boldsymbol{a}} + \dot{\mathbb{I}}^{\boldsymbol{b}} + \dot{\mathbb{I}}^{cross} \tag{5.24}$$

where the quantity

$$\dot{\mathbb{I}}^{cross} \coloneqq \sum_{a\geq a';b\geq b'} \jmath^{b,b'}_{a,a'} \log \frac{p(a',b')}{p(a,b)}. \tag{5.25}$$

This expression is the rate of Shannon entropy change of the joint system due to the cross-transitions. Using the non-negativity of $\Sigma^{\boldsymbol{a}}$ and $\Sigma^{\boldsymbol{b}}$, we arrive at the second law-like inequalities

$$\Sigma^{\boldsymbol{a}} = \sigma^{\boldsymbol{a}} - \dot{\mathbb{I}}^{\boldsymbol{a}} \geq 0, \tag{5.26}$$

$$\Sigma^{\boldsymbol{b}} = \sigma^{\boldsymbol{b}} - \dot{\mathbb{I}}^{\boldsymbol{b}} \geq 0. \tag{5.27}$$

where F^A and F^B are related by the change of mutual information and the rate of Shannon entropy in the cross-transitions according to equation (5.24).
The equations (5.26) and (5.27) are the same as previously derived for completely bipartite systems where two subsystems were coupled by regulatory transitions, but no cross-transitions were allowed [30]. In the absence of cross-transitions, the original framework [30] is recovered. Now, these inequalities have been generally derived for the systems where both regulatory and cross-transitions directly connecting the subsystems can take place. Such generalization is only possible if the definitions (5.21) and (5.22) are employed. Once the inequalities have been established, the same interpretation as in refs. [30] can be used.

Chapter 6

Stochastic thermodynamics & information processing of biochemical nanomachines

Generally systems may be classified by two different types of behaviour: Either they evolve towards maximum disorder or they spontaneously develop a high degree of organisation in space and time. The latter are example of dissipative systems at nonequilibrioum conditions, ranging from the Benard cell to biological living systems. living systems grow and develop with a constant supply of energy needed for reproduction and survival in varying conditions.

It is a number of coupled metabolic reactions and transport processes at the nanoscale that control the rate and timing of life processes. Biochemical processes maintain the biological cell in a nonequilibrium state by controlling the inflow of reactants and the outflow of products. Instead of decaying towards an equilibrium state, living biological systems increase in size and develop organised structures of high complexity. The core building blocks of cellular functions are the various interactions of macromolecules composed of proteins. For the synthesis and assembly of proteins and reproduction in general, cells convert energy in an efficient way for the transport of substrates across cell membranes and chemical gradients. Adenosine triphosphate (ATP) is the source of energy for cells, which is produced by oxidative phosphorylation in the inner membrane of the mitochondria.

One of the major challenges in modern biology is to understand how the molecular components of a living cell operate in a highly noisy environment.

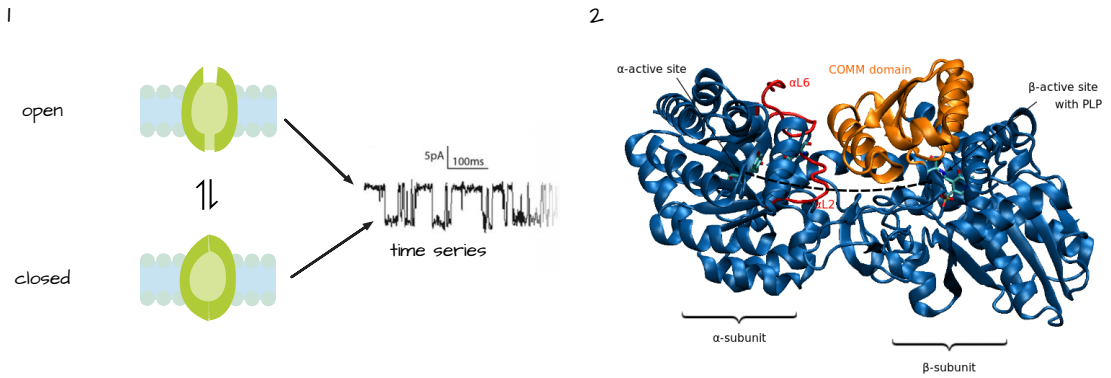

Figure 6.1: (1) Internal open and closed states of an ion channel. The time series of the opened and closed confirmation can be used to determine transition rates to formulate a Markov process. (2) Structure of the enzyme from *Salmonella Typhimurium* (PDB code: 2J9X [**?**]. The allosteric communication is mediated by the loops αL2 and αL6 at the α-site and the COMM domain at the β-site. The intramolecular tunnel is indicated by the dashed line.[10]

6.1 Internal and External States of a Biomolecule

Today it is still a big challenges is to understand how the biomolecular components of a living cell operate and interact in an environment of high thermal fluctuations. Macromolecules in a cell are subject to strong thermal fluctuations arising from the continual bombardment by molecules in the surrounding aqueous solution (cytosol). The magnitude of molecular fluctuations is set by the unit of thermal energy k_BT, the Boltzmann constant. At the molecular level it is useful to distinguish between *internal configurational states* and *external states* such as position and momentum of the molecule. The internal degrees of freedom can be represented as discrete states, e.g. configurational states. Due to the highly noisy environment, the transitions between these states are random and modelled as a continuous time Markov process as introduced above. For example a simple two-state continuous Markov process can be used to model the opening and closing of an ion channel. The transition rates can for instance be deduced from experiemnts, where the time series has been recorded (see figure 6.1).

6.2 The channeling enzyme tryptophan synthase

The enzyme tryptophan synthase (TS) yields a paradigmatic example of a chemical nanomachine. Several reviews devoted to tryptophan synthase are available [6, 19, 50, 17], with the authors charactering it as an allosteric molecular factory [6] and a channeling nanomachine [19]. TS is an enzyme that performs synthesis of the essential amino acid tryptophan from the substrates serine and indole glycerol phosphate (IGP). IGP is scarce inside the cell and therefore high catalytic efficiency is required. Furthermore, an intermediate product (indole) of the synthesis reaction is hydrophobic and can easily escape through the cell membrane. Therefore, its release into the cytoplasm must be avoided. As single molecule and with 13 different reaction steps catalyzed over a reaction cycle, TS is a biomolecule with a complex interplay of allosteric regulation to overcome the above mentioned critical constraints. During the biosynthesis cycle an intermediate product is channeled from one active center to another.

TS has been experimentally investigated extensively [35, 3, 66, 18, 11, 45]. Its mechanism is completely known, all intermediate chemical states and reaction steps have been identified and rate constants for practically all transitions have been measured. The allosteric regulatory interactions have been determined and quantified. Moreover, protein conformations corresponding to different states have been determined by using X-ray diffraction methods.

6.2.1 Tryptophan synthase as a chemical nanomachine

The molecular structure of tryptophan synthase is shown in Fig. S2. The protein consists of α and β-subunits, both with their own catalytic sites. The active centers of the subunits are connected by a 25 Å long intramolecular tunnel for transport of indole. Each catalytic site has its own gate controlling the release and uptake of substrates and products. The operation of tryptophan synthase involves a complex pattern of allosteric cross-regulation controlling the reactions and the configuration of the gates. It is known that the enzyme adopts two different conformational states: the catalytically inactive state with open gates ("open conformation") and the state with enhanced catalytic activity and closed gates ("closed conformation") depending on its chemical state CITES. The switching between the two conformational states is mediated by the COMM subdomain of the β-subunit.

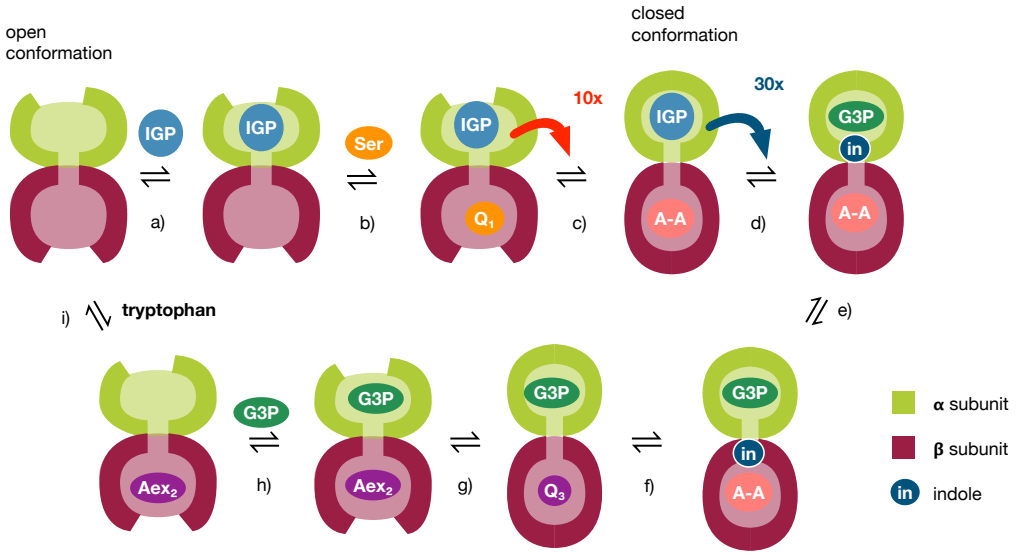

Figure 6.2: Schematic operation of tryptophan synthase. Aex_2 is the state of the β-subunit with tryptophan present inside it.

A simplified scheme of the catalytic cycle of tryptophan synthase with several omitted states is displayed in Fig.. Here, the α-subunit is shown in green and the β-subunit in blue. The catalytic cycle begins with the enzyme in the state where both sites are empty and the gates are open. Then, the substrate IGP binds to the α-subunit and serine to the β-subunit, where it is quickly converted to the serine quinoline intermediate (Q_1). IGP activates the formation of the α-aminoacrylate (A-A) and the enzyme adopts the closed conformation, as schematically shown in Fig. . In the state (IGP,A-A) where both gates are closed, A-A activates the cleavage of IGP to produce G3P and indole. Indole is then channeled to the β-site where it reacts with A-A to give the tryptophan quinoline intermediate (Q_3) that is converted to tryptophan (Aex_2 is the external aldimine of tryptophan in the β-subunit). In the state (G3P,Aex_2) the gates open and the products tryptophan and G3P are released. Thus the enzyme returns to the initial conformation (emtpy,empty) and is ready to start the next cycle. The functioning of tryptophan synthase can be also illustrated in a different way (Fig. 6.3). Each subunit undergoes stochastic transitions that either represent internal chemical transformations or correspond to binding and release of substrates and products. Both subunits are coupled to chemical reservoirs where constant substrate and product concentrations are maintained. These reservoirs can also be considered as *chemostats*. Generally, there is a difference of chemical potentials

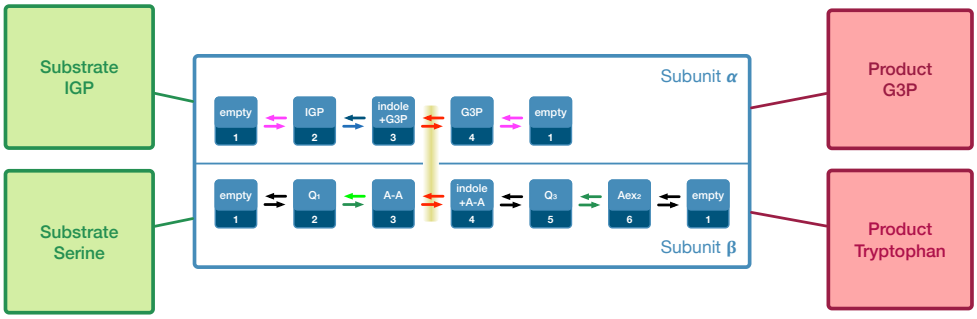

Figure 6.3: Allosteric cross-regulation and channeling in tryptophan synthase. Magenta: transitions blocked in the states A-A, A-A + indole and Q_3 of the β-site. Green (light and dark): blocked in the state empty of the α-site. Light green: enhanced by a factor of 9.7 in the state IGP of the α-site. Blue (light and dark): blocked in the states empty, Q_1, Aex_2 of the β-site. Light blue: enhanced by a factor of 27.7 in the state A-A of the β-site. Red: Channeling instantaneously changes the states of both sites. Numbers represent numerical notations of the respective chemical states of a subunit. Aex_2 is the state of subunit β with tryptophan inside it.

α-site	empty	IGP	indole+G3P	G3P		
Variable a	1	2	3	4		
β-site	empty	Q_1	A-A	indole+A-A	Q_3	Aex_2
Variable b	1	2	3	4	5	6

Table 6.1: Enumeration of chemical states of α- and β-subunits by variables a and b.

between the substrate and the product chemostats. Because of such a difference, the enzyme is out of equilibrium and can operate as a chemical nanomachine. There is an extensive pattern of allosteric cross-regulation between the two subunits. The transitions empty $\rightleftharpoons$ IGP and G3P $\rightleftharpoons$ empty (magenta) in the α-site are blocked (i.e., the gate in the α-subunit is closed) in the states A-A, A-A + indole and Q_3 of the β-site. The transitions IGP $\rightleftharpoons$ indole+G3P (light and dark blue) in the α-site are blocked in the states empty, Q_1, Aex_2 of the β-site. The rate of the transition IGP $\rightarrow$ indole+G3P (light blue) in the α-site is enhanced by a factor of 27.7 in the state A-A of the β-site. The transitions $Q_1 \rightleftharpoons$ A-A and $Q_3 \rightleftharpoons Aex_2$ (green) in the β-site are blocked in the state empty of the α-site. The transition $Q_1 \rightarrow$ A-A (light green) in the β-site is enhanced by a factor of 9.7 in the state IGP of the α-site. The changes indole+G3P $\rightleftharpoons$ G3P and A-A $\rightleftharpoons$ indole+A-A (red) corresponding to indole channeling from the α- to the β-site occur simultaneously and represent a single stochastic transition.

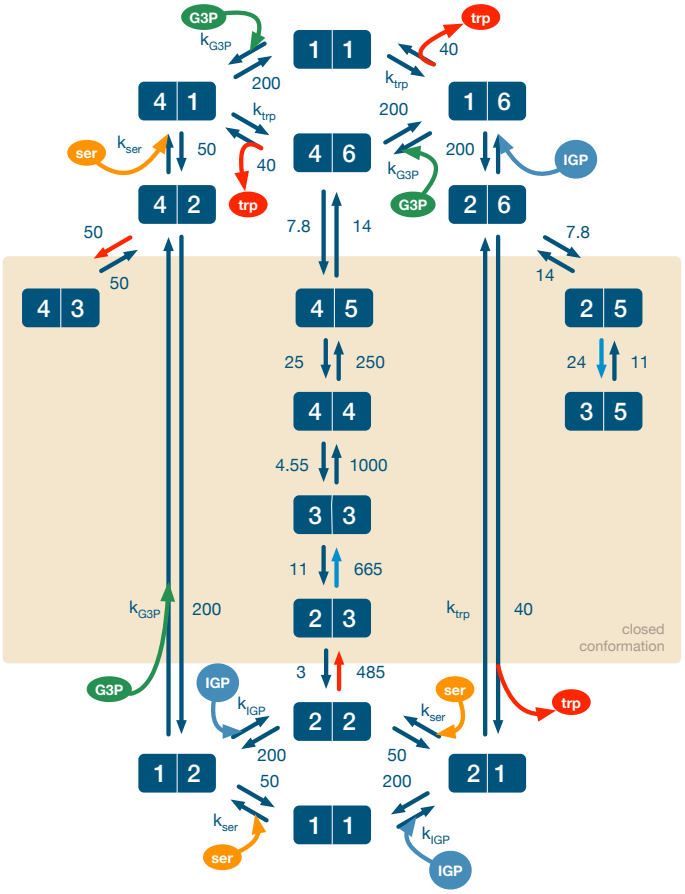

Figure 6.4: The kinetic Markov network of tryptophan synthase with numerical values of all transition rates in units of s^{-1}.

To simplify the notations, the states of both subunits will be enumerated below. The assignment of integer numerical variables a and b to different states of the α- and β-subunits is given in Table 6.1 and introduced in Fig. 6.3.

6.2.2 The Kinetic Markov Network

The stochastic single-molecule operation of tryptophan synthase can be seen as random wandering over a Markov network of its internal states. In this network, each binary internal state (a, b) with $a = 1, ..., 4$ and $b = 1, ..., 6$ represents a possible combination of the individual states a and b of the α- and the β-subunits. The network is shown in Fig. 6.4. The nodes of the network are different chemical states and arrows indicate the transitions between them. Additionally, ligand binding and release are displayed. The states within the colored box correspond to the closed conformation

Table 6.2: Ligand binding rate constants k, ligand concentrations c and the respective transition rates $w = kc$ under physiological conditions.

Reaction	Binding constant k	Concentration c	Rate w	Reference
β-empty + Ser $\rightarrow$ Q_1	$7.5 \cdot 10^{-2}$ $\mu M^{-1}s^{-1}$	c(Ser)=68 μM	5.1 s^{-1}	[40]
β-empty + Trp $\rightarrow$ Aex_2	0.15 $\mu M^{-1}s^{-1}$	c(Trp)=12 μM	1.8 s^{-1}	[40]
α-empty + IGP $\rightarrow$ α-IGP	10 $\mu M^{-1}s^{-1}$	c(IGP)=3.5μM	35 s^{-1}	[2, 39]
α-empty + G3P $\rightarrow$ α-G3P	0.2 $\mu M^{-1}s^{-1}$	c(G3P)=49 μM	9.8 s^{-1}	[2]

of the enzyme where the gates are closed and the ligands cannot arrive or be released. The numbers next to the arrows give the respective transition rates in units of s^{-1}. Note that the bottom and upper states $(1, 1)$ in Fig. 6.4 are identical; they are shown separately only for convenience in the displayed network.

The kinetic Markov network for tryptophan synthase has been constructed in the previous publication [43] where all transition rates have been determined from the available experimental data. In that publication, we have modeled, however, a typical experimental situation where product concentrations remain vanishingly small. Therefore, product binding events were not included into the scheme. Moreover, the reverse indole channeling transitions were not taken into account, because such reaction events have never been experimentally observed. Additionally, the reverse reaction in the beta-subunit was not included. Because of this, the previously constructed network has been partially irreversible.

In contrast to this, the network shown in Fig. 6.4 is fully reversible. The enzyme is now assumed to operate under typical physiological conditions where products are present in substantial concentrations and product binding events can therefore take place. Moreover, as shown below, we can use additional experimental data and thermodynamic consistency conditions to determine the missing rate of reverse indole channeling. Additionally, the rate for the reverse reaction $Q_3 \rightarrow$ indole+A-A in the β-subunit has been taken from [41]. Numerical values of the rate constants of each transition are shown next to the respective arrows in Fig. 6.4. The rates of the binding of substrate and product molecules are proportional to substrate and product concentrations. The substrate and product concentrations under physiological conditions were taken from [7]. The respective binding rate constants are given in Table 6.2.

For the combined states (a, b), time dependent probabilities $p(a, b; t)$ can be intro-

duced. They satisfy the master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}p(a,b;t) = \sum_{a'=1}^{4}\sum_{b'=1}^{6}[W_{a,a'}^{b,b'}p(a',b';t) - W_{a',a}^{b',b}p(a,b;t)] \tag{6.1}$$

where $W_{a,a'}^{b,b'}$ denotes the transition rate from a state (a',b') to the state (a,b). The numerical values of the transition rates are shown in Fig. 6.4

As we have already noticed (see, e.g., Fig. 6.2), only the transition representing indole channeling involves simultaneous changes of the states of both α- and β-sites. All other transitions change the state of only one subunit although the rates of such transitions can be controlled by the state of the other subunit. Therefore, the Markov network of tryptophan synthase has a special structure. It is almost bipartite (see [28, 30, 16]) and the transition matrix elements can be written as

$$w_{a,a'}^{b,b'} = \begin{cases} W_{a}^{b,b'} & \text{if } a = a' \\ W_{a,a'}^{b} & \text{if } b = b' \\ W_{4,3}^{4,3} & \text{if } (a',b') = (3,3) \text{ and } (a,b) = (4,4) \\ W_{3,4}^{3,4} & \text{if } (a',b') = (4,4) \text{ and } (a,b) = (3,3) \\ 0 & \text{else.} \end{cases} \tag{6.2}$$

The indole channeling couples the two subunits and perturbs the complete bipartite structure of the Markov network.

It is again convenient to write the master equation as a continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}t}p(a,b;t) = \sum_{a',b'} j_{a,a'}^{b,b'}. \tag{6.3}$$

Taking into account the special form (6.2) of the transition matrix, the master equation can also be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}p(a,b;t) = \sum_{b'} j_{a}^{b,b'} + \sum_{a'} j_{a,b'}^{b} + [\delta_{(a,b)}^{(3,3)} - \delta_{(a,b)}^{(4,4)}]j^{channel} \tag{6.4}$$

where $\delta_x^y = 1$, if $x = y$ and $\delta_x^y = 0$ otherwise. The currents corresponding to transitions inside the β-subunit are $j_a^{b,b'} = W_a^{b,b'}p(a,b';t) - W_a^{b',b}p(a,b;t)$ and the currents $j_{a,a'}^b$ for the transitions within the α-subunit are defined similarly. The flux corresponding to channeling is $j^{channel} = W_{4,3}^{4,3}p(3,3;t) - W_{3,4}^{3,4}p(4,4;t)$.

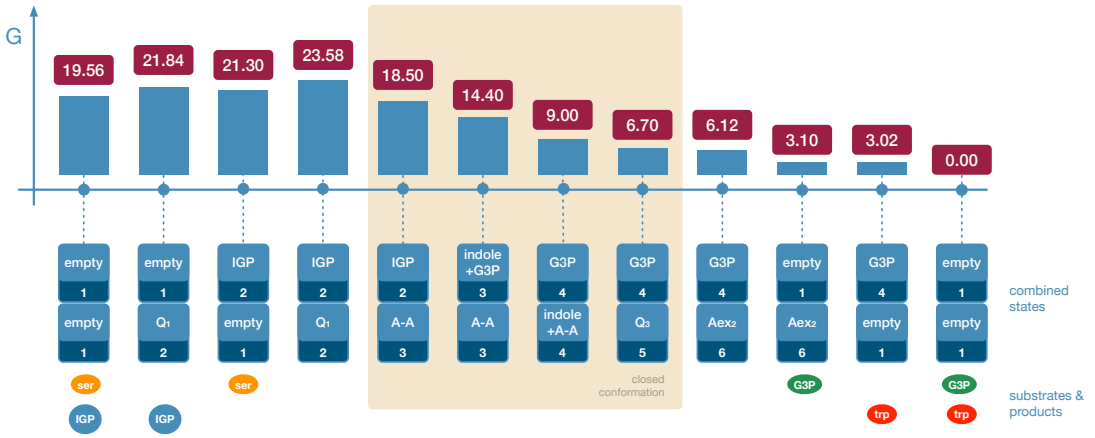

Figure 6.5: The Gibbs energy landscape along the main pathway of tryptophan synthase under physiological ligand concentrations. The Gibbs energies are given in units of k_BT. Physiological concentration values are chosen. In the states within the beige box, the molecular gates are closed and the enzyme is disconnected from the chemostats. Tryptophan is present inside the β-subunit in the state Aex_2.

6.2.3 Energies and Detailed Balance

Thermodynamics implies that all transitions between the states should satisfy the condition of detailed balance. This condition is that at thermal equilibrium the net probability flux between any two states is absent. For the considered network it implies that the ratio of the rates $W^{b',b}_{a',a}$ and $W^{b,b'}_{a,a'}$ for forward and backward transitions between any two states (a,b) and (a',b') satisfies the equation

$$\frac{W^{b,b'}_{a,a'}}{W^{b',b}_{a',a}} = \exp\left(\frac{G(a,b) - G(a',b')}{k_BT}\right) \tag{6.5}$$

where $G(a,b)$ and $G(a',b')$ are Gibbs energies of the respective states in the network at equilibrium, T is the temperature, and k_B is the Boltzmann constant.
For transitions between the states $(a,b) \to (a',b')$ that do not involve binding or release of ligands, the rates $W^{b,b'}_{a,a'}$ coincide with the respective rate constants $k^{b,b'}_{a,a'}$ and the Gibbs energies $G(a,b)$ are the internal Gibbs energies $g(a,b)$ of the molecular states. In this case, equation (6.5) takes the form

$$\frac{k^{b,b'}_{a,a'}}{k^{b',b}_{a',a}} = \exp\left(\frac{g(a,b) - g(a',b')}{k_BT}\right). \tag{6.6}$$

Note that, for macromolecules the Gibbs energies $g(a,b)$ of internal states are different from the internal energies $\epsilon(a,b)$ of such states because they additionally include

entropic contributions and solvent effects.

The transitions that involve binding or release of a ligand should be treated separately. Suppose that a transition from (a, b) to (a', b') is accompanied by binding of a ligand and the ligand is released in the backward transition. Then the forward transition rate is proportional to the ligand concentration c, i.e. $W^{b',b}_{a',a} = k^{b',b}_{a',a}c$, whereas for the backward transition we have $W^{b,b'}_{a,a'} = k^{b,b'}_{a,a'}$. Moreover, the Gibbs energies in (6.5) include now contributions from ligand particles, i.e. $G(a, b) = g(a, b) + \mu$, where μ is the chemical potential of the ligand. For the considered weak solutions, we have $\mu = \mu_0 + k_B T \ln c$. Substitution of these expressions into equation (6.5) yields

$$\frac{k^{b,b'}_{a,a'}}{k^{b',b}_{a',a}} = \exp\left(\frac{g(a, b) + \mu_0 - g(a', b')}{k_B T}\right). \tag{6.7}$$

In this equation, the ligand can be either a substrate or a product if reverse binding of a product molecule takes place.

As shown by Schnakenberg [51], one can derive further identities by considering different pathways in a Markov network. Suppose that the chosen pathway represents a closed cycle Γ that involves only the internal states of the molecule without the events of ligand binding or release. Then, by using equation (6.6), one can show that the identity

$$\prod_{\Gamma} \frac{k^{b,b'}_{a,a'}}{k^{b',b}_{a',a}} = \exp\left(\sum_{\Gamma} \frac{g(a, b) - g(a', b')}{k_B T}\right) = 1 \tag{6.8}$$

holds, with the multiplication on the left side performed over all transitions that belong to the chosen cycle.

If the pathway Γ involves a conversion of substrate to a product or back, application of condition (6.7) leads to a modified identity. For tryptophan synthase, it has the form

$$\prod_{\Gamma} \frac{W^{b,b'}_{a,a'}}{W^{b',b}_{a',a}} = \exp\left(\frac{\mu(\mathrm{trp}) + \mu(\mathrm{G3P}) - \mu(\mathrm{ser}) - \mu(\mathrm{IGP})}{k_B T}\right) \tag{6.9}$$

if the pathway Γ leads from the bottom to the top empty states $(1, 1)$ in the Markov network in Fig. 6.4, i.e. if it corresponds to conversion of the two substrate molecules IGP and serine to the two product molecules G3P and tryptophan.

The detailed balance condition (6.5) and the Schnakenberg identities (6.8) and (6.9) can be used to check the thermodynamic consistency of a Markov network, to find

missing rate constants of some transitions, and to determine Gibbs energies of different states. Particularly, in the Markov network of tryptophan synthase, there is a transition from the state $(4,4)$ to $(3,3)$ that corresponds to the channeling of indole from the β- to the α-site. This transition has never been observed experimentally and its rate constant could not be measured. This rate constant can however be determined, as explained below, by using the identity (6.9) and additional experimental data.

In each turnover cycle of tryptophan synthase two substrate molecules (IGP and serine) are converted into two product molecules (G3P and tryptophan) and some energy Δq is dissipated during this conversion process. In microcalorimetric measurements under isothermal conditions, it is possible to determine the heat Q released by the reaction in a given volume per unit time. This heat can be estimated as $Q = N\Delta q/\tau$ where τ is the mean turnover time and N is the number of enzyme molecules in the reaction volume. The last two quantities can be determined independently and thus the amount of heat Δq in one turnover cycle can be obtained. Such calorimetric measurements have been performed for tryptophan synthase, under standard concentration conditions $c_0(\mathrm{IGP}) = c_0(\mathrm{ser}) = c_0(\mathrm{G3P}) = c_0(\mathrm{trp}) = 1$ M, and they yield $\Delta q = 20.46\ k_BT$ [36].

On the other hand, the released heat Δq corresponds to the difference of chemical potentials of substrates and products

$$\Delta q = \mu(\mathrm{IGP}) + \mu(\mathrm{ser}) - \mu(\mathrm{G3P}) - \mu(\mathrm{trp}) \tag{6.10}$$

where, for example, $\mu(\mathrm{IGP}) = \mu_0(\mathrm{IGP}) + k_BT \ln c(\mathrm{IGP})$. Thus, the difference of the chemical potentials $\mu_0(\mathrm{IGP}) + \mu_0(\mathrm{ser}) - \mu_0(\mathrm{G3P}) - \mu_0(\mathrm{trp})$ in tryptophan synthase under standard conditions is equal to $20.46\ k_BT$.

By using the identity (6.9) and the known value of Δq for tryptophan synthase, reverse channeling transition rate can be determined as $k_{3,4}^{3,4} = 4.55\ \mathrm{s}^{-1}$. This is indeed much smaller than the measured rate $k_{4,3}^{4,3} = 1000\ \mathrm{s}^{-1}$ of the forward channeling transition. Therefore, the reverse channeling transitions should be very rare and this is why they have not been experimentally observed.

Furthermore, the detailed balance conditions (6.5) and (6.9) can be used to determine, by repeated application, Gibbs energies $G(a,b)$ with respect to the Gibbs energy of a certain reference state.

Our reference state corresponds to the free enzyme with two products (tryptophan

and G3P) and its Gibbs energy is chosen as $G_{final} = 0$. In the initial state, the enzyme is free, there are two additional substrate molecules (serine and IGP) and the two product molecules (tryptophan and G3P) are missing. The Gibbs energy of the initial state is therefore $G_{initial} = \mu(\text{IGP}) + \mu(\text{ser}) - \mu(\text{G3P}) - \mu(\text{trp})$. It should be noted that it depends on the involved ligand concentrations c because $\mu = \mu_0 + k_B T \ln c$. It coincides with the amount of heat Δq released in one turnover cycle. The value above given $\Delta q = 20.46\ k_B T$ corresponds to the standard conditions $c_0(\text{IGP}) = c_0(\text{ser}) = c_0(\text{G3P}) = c_0(\text{trp}) = 1$ M. Recalculating this under the physiological concentrations (Table 6.2), we obtain $G_{initial} = \Delta q = 19.56\ k_B T$.

There are also several states where one of the subunits is empty and the other subunit has a ligand bound to it. For example, the state (IGP, empty) has IGP bound to the α-subunit and no ligand in the β-subunit. The Gibbs energy of this state is $G(\text{IGP,empty}) = g(\text{IGP,empty}) - g_0 + \mu(\text{ser}) - \mu(\text{G3P}) - \mu(\text{trp})$. It includes both the difference of the chemical potentials, depending on the concentrations, and the internal Gibbs energies $g(\text{IGP,empty})$ and $g_0 = g(\text{empty,empty})$ of the state (IGP, empty) and the free state of the enzyme.

Finally, there are states where both subunits are occupied. For example, for the state (IGP, Q_1), we have $G(\text{IGP,}Q_1) = g(\text{IGP,}Q_1) - g_0 - \mu(\text{G3P}) - \mu(\text{trp})$. For the state (IGP, A-A), we have $G(\text{IGP,A-A}) = g(\text{IGP,A-A}) - g_0 - \mu(\text{G3P}) - \mu(\text{trp})$. Note that the difference $G(\text{IGP,}Q_1) - G(\text{IGP,A-A}) = g(\text{IGP,}Q_1) - g(\text{IGP,A-A})$ is determined only by the internal Gibbs energies of the states and is independent of ligand concentrations. This difference gives the amount of heat dissipated in the respective transition.

Figure 6.5 shows the Gibbs energy landscape of tryptophan synthase along its main pathway. After the binding of substrates requiring activation energies of 1.74 $k_B T$ for IGP binding and 2.28 $k_B T$ for serine binding, all transitions towards product formation are exergonic. The four catalytically important transitions (IGP,Q_1) $\leftrightarrow$ (IGP,A-A) $\leftrightarrow$ (indole+G3P,A-A) $\leftrightarrow$ (G3P,indole+A-A) $\leftrightarrow$ (G3P,Q_3) in the closed conformation of the enzyme are highly exergonic and accompanied by heat release in the range between 5.40 and 2.30 $k_B T$. The release of the products G3P and tryptophan is accompanied by the heat release of 3.10 and 3.02 $k_B T$, respectively.

6.2.4 Entropy Production and Flow

The Shannon entropy at time t is given by

$$\mathbb{H}(t) = -\sum_{a,b} p(a,b;t) \ln p(a,b;t) \tag{6.11}$$

where $p(a,b;t)$ is the probability to find the enzyme in the state (a,b) at time t. Its rate of change is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{H} = \frac{1}{2} \sum_{a,a',b,b'} j^{b,b'}_{a,a'} \log \frac{p(a',b';t)}{p(a,b;t)} \tag{6.12}$$

This rate of change can be decomposed as

$$d_t\mathbb{H} = \sigma - h \tag{6.13}$$

into the difference of the entropy production σ inside the enzyme and of the net flow h of entropy *from* the enzyme, i.e. of the rate of entropy export by it.
According to equation (6.12), the rate of change of the total entropy of the enzyme can be also written as a sum of the rates of entropy change $s^{b,b'}_{a,a'}$ in each individual transition, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{H} = \frac{1}{2} \sum_{a,a',b,b'} s^{b,b'}_{a,a'}, \text{ with } s^{b,b'}_{a,a'} = j^{b,b'}_{a,a'} \log \frac{p(a',b')}{p(a,b)}. \tag{6.14}$$

The same holds for the total entropy production σ and the rate of entropy export h. Thus the quantities $\sigma^{b,b'}_{a,a'}$ and $h^{b,b'}_{a,a'}$ corresponding to individual transitions can be introduced,

$$h = \frac{1}{2} \sum_{a,a',b,b'} h^{b,b'}_{a,a'};\ \sigma = \frac{1}{2} \sum_{a,a',b,b'} \sigma^{b,b'}_{a,a'} \tag{6.15}$$

where

$$h^{b,b'}_{a,a'} = J^{b,b'}_{a,a'} \log \frac{w^{b,b'}_{a,a'}}{w^{b',b}_{a',a}}, \tag{6.16}$$

$$\sigma^{b,b'}_{a,a'} = J^{b,b'}_{a,a'} \log \frac{w^{b,b'}_{a,a'} p(a',b')}{w^{b',b}_{a',a} p(a,b)}. \tag{6.17}$$

Note that $\sigma^{b,b'}_{a,a'} = \sigma^{b',b}_{a',a}$, $h^{b,b'}_{a,a'} = h^{b',b}_{a',a}$ and $s^{b,b'}_{a,a'} = s^{b',b}_{a',a}$ and therefore these quantities do not depend on the choice of the transition direction, i.e. they are the same

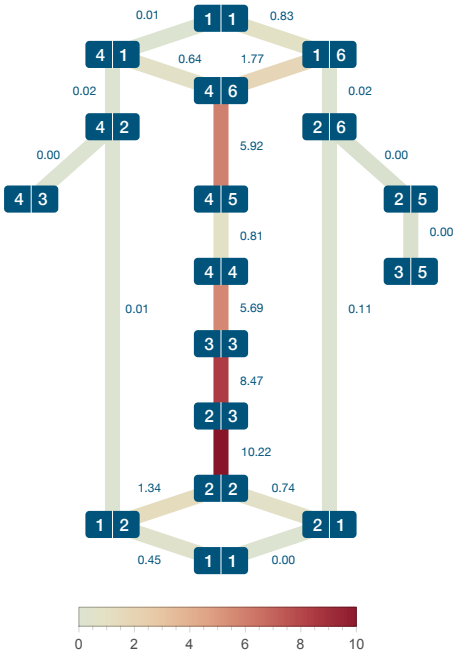

Figure 6.6: Entropy production in different transitions in the nonequilibrium steady-state. The values of entropy production are given in units of bit s^{-1} next to the links between the states. Additionally, color coding of the links according to the corresponding entropy production is used.

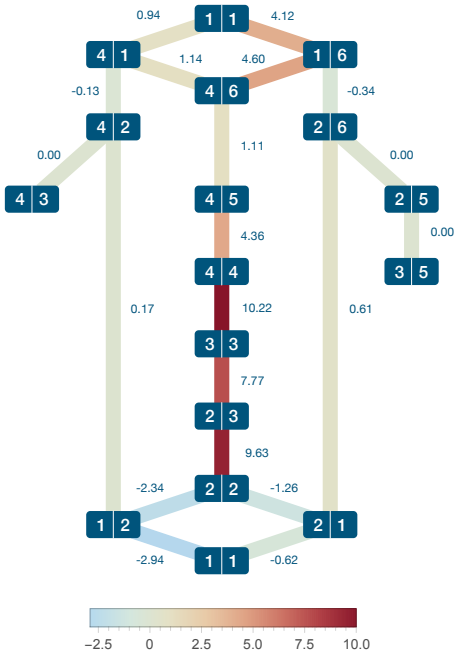

Figure 6.7: Rates of entropy export in individual transitions in tryptophan synthase. The same notations as in Fig. 6.6.

both for $(a, b) \to (a', b')$ and $(a', b') \to (a, b)$. The entropy production $\sigma_{a,a'}^{b,b'}$ characterizes dissipation or irreversibility whereas the entropy flow $h_{a,a'}^{b,b'}$ represents the average heat exchanged per unit time with the environment in a transition. We have $s_{a,a'}^{b,b'} = \sigma_{a,a'}^{b,b'} - h_{a,a'}^{b,b'}$.

In the state of thermal equilibrium, all fluxes $j_{a,a'}^{b,b'}$ vanish and therefore according to equations (6.16) and (6.17) there are no transitions where entropy is produced or exported. Under physiological conditions, however, the enzyme tryptophan synthase operates far from thermal equilibrium, with the difference of Gibbs energies of 19.56 $k_B T$ for one cycle. Thus, its operation is characterized by nonequilibrium steady-state. In the respective nonequilibrium steady-state with the stationary probability distribution $\bar{p}(a, b)$, the fluxes $\bar{j}_{a,a'}^{b,b'}$ do not vanish and therefore the transitions are accompanied by entropy production and entropy export. Because the entropy $\mathbb{H}$

is conserved in this state, $d_t\mathbb{H} = \sigma - h = 0$. Hence the total entropy production σ is counterbalanced by the entropy export h. Note that, although $d_t\mathbb{H} = 0$, the rates of entropy change $s_{a,a'}^{b,b'}$ for individual transitions are not zero even in the nonequilibrium steady-state.

The stationary probability distribution $\bar{p}(a, b)$ can be found by solving the master equation (6.1) in the nonequilibrium steady-state. Numerical values of the probabilities $\bar{p}(a, b)$ corresponding to all possible states are given Table S1. Then, by using equation (**??**), the fluxes $\bar{J}_{a,a'}^{b,b'}$ can be determined. According to equations (6.14), (6.16) and (6.17), the values of $\sigma_{a,a'}^{b,b'}$, $h_{a,a'}^{b,b'}$ and $s_{a,a'}^{b,b'}$ can be calculated afterwards.

The results are displayed in Figs. 6.6 and 6.7. Here, we show the same network as in Fig. 6.4, but, for simplicity, retain only numerical notations of the states. Only the links between the states are shown because the transition directions are not important (see the comment after equation (6.17)) For each link, the value of the quantities $\sigma_{a,a'}^{b,b'}$ or $h_{a,a'}^{b,b'}$ is indicated. Additionally, color coding is used.

Here and below, all numerical values for entropy and information are given in units of bits. We have 1 bit $= \ln 2 = 0.693$, because natural logarithms are used in our definition of the Shannon entropy.

The rates of entropy or information change are given in bits per seconds. Alternatively, they can also be expressed by the respective amounts per a catalytic cycle. Note that the substrate conversion rate of the enzyme is equal to the probability flux $J^{channel}$ because each productive cycle includes this transition. The mean catalytic cycle time is the inverse of the substrate conversion rate. Under physiological concentrations, we find that the mean cycle time is 0.75 s. Tryptophan synthase is a slow molecular machine.

Figure 6.6 shows numerical values of entropy production for all individual transitions within the enzyme. The entropy is mostly produced along the main catalytic pathway. The highest entropy production (10.22 bit s^{-1}) is found for the allosterically activated transition $Q_1 \leftrightarrow$ A-A in the β-site. In contrast to this, all transitions involving futile states (side branches of the network) have values of entropy production below 0.01 bit s^{-1} per second. Ligand binding and release is characterized by entropy production below 1.78 bit s^{-1} per second.

The values for entropy export are given in Fig. 6.7. The entropy export takes is maximal (between 7.77 bit s^{-1} and 10.22 bit s^{-1}) for the transitions (IGP,Q_1) $\leftrightarrow$

(IGP,A-A) $\leftrightarrow$ (indole+G3P,A-A) $\leftrightarrow$ (G3P,indole+A-A) where most of the heat exchange with the environment takes place. All other transitions have absolute values smaller than 4.61 bit s^{-1}. Note that transition (G3P,Q_3) $\leftrightarrow$ (G3P,Aex_2) has a small entropy export, but a high entropy production.

Because the rate of entropy change in a transition is given by the difference of entropy production and export, this rate can be found by subtracting the respective values in Figs. 6.6 and 6.7. Thus the transition (G3P,Q_3) $\leftrightarrow$ (G3P,Aex_2) in the main catalytic pathway has the largest rate of entropy increase $s_{4,4}^{6,5} = 4.82$ bit s^{-1}. In contrast to this, channeling and the subsequent transition (G3P,indole+A-A) $\leftrightarrow$ (G3P,Q_3) are accompanied by the net export of entropy at the rates $s_{3,3}^{4,4} = -4.53$ bit s^{-1} and $s_{4,4}^{5,4} = -3.55$ bit s^{-1}.

Using the computed rates of entropy production and export for individual transitions, total amounts for the whole enzyme per a turnover cycle can be obtained. We find that, within a single catalytic cycle of tryptophan synthase, 27.79 bits of entropy are produced. The same amount of entropy is on the average exported by the enzyme per one cycle.

6.2.5 Information Exchange between the Subunits

There is a complex pattern of allosteric interactions between the two subunits of tryptophan synthase. Additionally, one transition that corresponds to indole channeling and affects simultaneously both subunits takes place. The allosteric cross-regulations and channeling lead to the development of correlations between the internal states of the subunits. Previously, the presence of correlations has been demonstrated by computing the Pearson correlation coefficients for all possible pairs of states [43]. In this section, the concept of mutual information will be employed to further quantify the effects of allosteric cross-regulation and channeling. Our method is based on the theory of information exchange between the subsystems [28, 30, 16] which had, however, to be extended to the situation where transitions involving simultaneously both subunits can also occur. The stochastic mutual information $i(a, b)$ of a pair of states (a, b) of the two subunits is defined as

$$i(a,b) = \ln \frac{p(a,b)}{p_\alpha(a) p_\beta(b)} \tag{6.18}$$

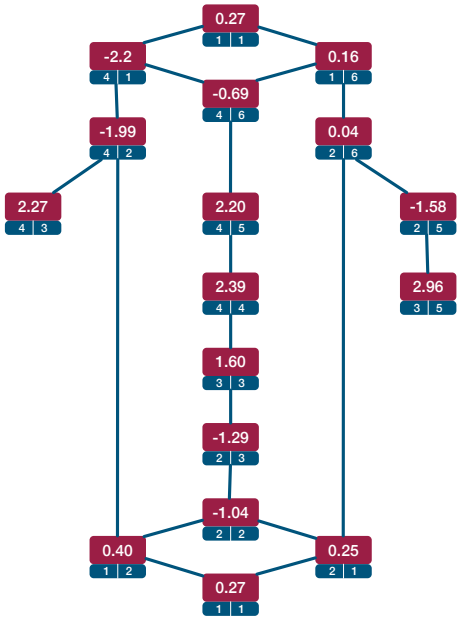

Figure 6.8: Stochastic mutual information $i(a,b)$ in units of bit for different states.

where $p_\alpha(a) = \sum_{b=1}^{6} p(a,b)$ and $p_\beta(b) = \sum_{a=1}^{4} p(a,b)$ are the probability distributions for the states of α- and β-subunits. The stochastic mutual information $i(a,b)$ quantifies correlations between the states a of the α-subunit and b of the β-subunit, it vanishes if these states are statistically independent, i.e. if $p(a,b) = p_\alpha(a)p_\beta(b)$. If the stochastic mutual information is negative, anti-correlations between the stats are present.

The values $i(a,b)$ under physiological conditions are shown in Fig. 6.8 for all states (a,b). We find high positive mutual information (2.39 and 2.20 bits) in the states (G3P,indole+A-A) and (G3P,Q_3) after indole channeling and after the indole reaction at the β-site in the main pathway. This agrees with the previous analysis using the Pearson correlation coefficients [43]. As a result of channeling, both subunits simultaneously arrive at the state (G3P,indole+A-A) and high positive correlations are characteristic for it. On the other hand, mutual information is negative (-1.04 bits) in the state (IGP,A-A) before channeling. This is an effect of allosteric interactions: when the β-subunit is in the state A-A, the cleavage of IGP into G3P and indole is blocked when the β-subunit is in the state Q_1, but it is possible in the state A-A. The statistical average of $i(a,b)$ over all pair states (a,b) yields the mutual information I of the whole system

$$\mathbb{I} = \sum_{a=1}^{4}\sum_{b=1}^{6} p(a,b)\ln\frac{p(a,b)}{p_\alpha(a)p_\beta(b)} = \sum_{a=1}^{4}\sum_{b=1}^{6} p(a,b)i(a,b). \tag{6.19}$$

This property is positive and it characterizes the strength of statistical correlations between the α- and β-subunits. For tryptophan synthase under physiological conditions we have $I = 0.49$ bit.

As shown in the Appendix, the rate of mutual information change for the entire system is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{I} = \frac{1}{2}\sum_{a,a'=1}^{4}{}' \iota^{\alpha}_{a,a'} + \frac{1}{2}\sum_{b,b'=1}^{6}{}' \iota^{\beta}_{b,b'} + \iota^{channel}. \tag{6.20}$$

Here, the sums exclude the forward and backward channeling transitions and we have

$$\iota^{\alpha}_{a,a'} = \sum_{b} j^{b}_{a,a'}\left[i(a,b) - i(a',b)\right], \tag{6.21}$$

$$\iota^{\beta}_{b,b'} = \sum_{a} j^{b,b'}_{a}\left[i(a,b) - i(a,b')\right], \tag{6.22}$$

$$\iota^{channel} = j^{channel}\left[i(4,4) - i(3,3)\right]. \tag{6.23}$$

Note that in a steady state $d_t\mathbb{I} = 0$ and therefore the terms (6.21) - (6.23) satisfy one additional constraint. Moreover, the terms $\iota^{\alpha}_{a,a'}$ and $\iota^{\beta}_{b,b'}$ do not depend on the choice of a direction for the transitions between a and a' or between b and b'. The quantity $\iota^{\alpha}_{a,a'}$ gives the contribution by the transition between the states a and a' in the α-subunit to the rate of change of the total mutual information of the system; this contribution is averaged over all possible regulatory states of the subunit β. A similar interpretation holds for the quantity $\iota^{\beta}_{b,b'}$.

By solving the master equation under physiological concentrations, we obtain the steady state probabilities $\bar{p}(\alpha,\beta)$. Substituting them into the equations (6.21) - (6.23) yields the values for $\iota^{\alpha}_{a,a'}$, $\iota^{\beta}_{b,b'}$ and $\iota^{channel}$. Figure 6.9 shows how the generation (or loss) of mutual information is distributed over the network. Mutual information is generated in three transitions in the α-subunit. Its highest generation rate is 3.79 bit s^{-1} in the transition (IGP $\leftrightarrow$ indole+G3P) preceding channeling. The channeling transition itself generates mutual information at a smaller rate (1.04 bit s^{-1}). All transitions in the β-subunit are accompanied by mutual information loss with the highest rate (-3.79 bit s^{-1}) achieved in the transition immediately after channeling ($\mathrm{Q}_3 \leftrightarrow \mathrm{Aex}_2$).

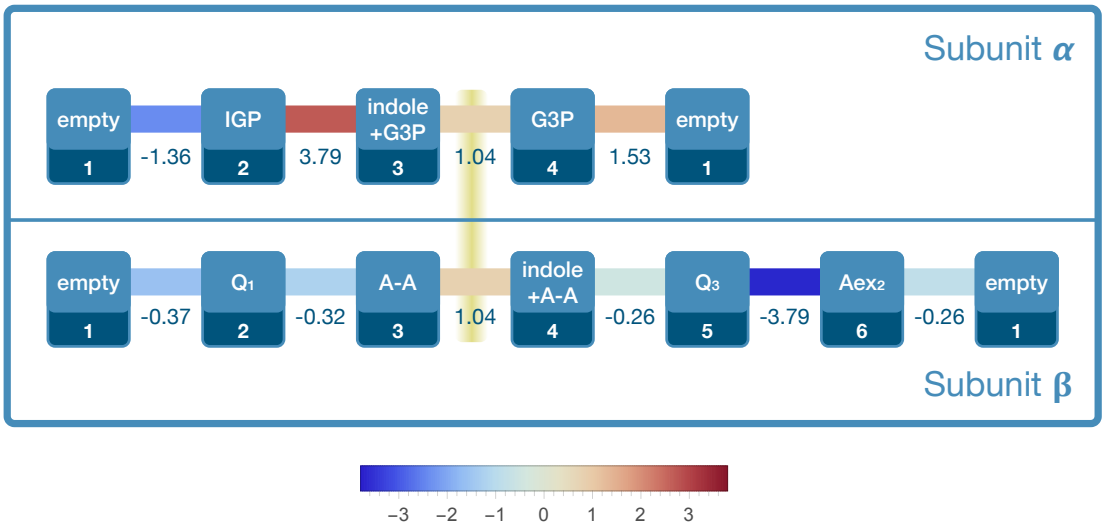

Figure 6.9: Rates of change of mutual information in units of bits per second for the transitions within α- and β-subunits and for the channeling transition.

Furthermore, information interactions between entire subunits can also be discussed. To do this, we write the rate of change of mutual information in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{I} = \dot{\mathbb{I}}^{\alpha} + \dot{\mathbb{I}}^{\beta} + \dot{\mathbb{I}}^{channel} \tag{6.24}$$

where

$$\dot{\mathbb{I}}^{\alpha} = \frac{1}{2}\sum_{a,a',b} \iota^{\alpha}_{a,a'} + \iota^{\alpha}_{channel}, \tag{6.25}$$

$$\dot{\mathbb{I}}^{\beta} = \frac{1}{2}\sum_{a,b,b'} \iota^{\beta}_{b,b'} + \iota^{\beta}_{channel}. \tag{6.26}$$

Here, we have divided the rate of generation of mutual information in the channel $\iota^{channel}$, given by equation (6.23) into three parts, i.e. $\iota^{channel} = \dot{\mathbb{I}}^{channel} + \iota^{\alpha}_{channel} + \iota^{\beta}_{channel}$, where

$$\dot{\mathbb{I}}^{channel} = j^{channel} \log \frac{p(3,3)}{p(4,4)}, \tag{6.27}$$

$$\iota^{\alpha}_{channel} = j^{channel} \log \frac{p_{\alpha}(3)p(4,4)}{p_{\alpha}(4)p(3,3)}, \tag{6.28}$$

$$\iota^{\beta}_{channel} = j^{channel} \log \frac{p_{\beta}(3)p(4,4)}{p_{\beta}(4)p(3,3)}. \tag{6.29}$$

Thus the rates of mutual information change in α- and β-subunits include now contributions $\iota^{\alpha}_{channel}$ and $\iota^{\beta}_{channel}$ from the channeling transition. The advantage of this definition is that, as shown in the Appendix, important thermodynamic inequalities for entropy production in both subunits become then satisfied.

In a steady state, $d_t\mathbb{I}$ vanishes and we have $\dot{\mathbb{I}}^{\alpha} + \dot{\mathbb{I}}^{\beta} + F^{channel} = 0$. If the channeling were absent, we would have had $\dot{\mathbb{I}}^{\alpha} = -\dot{\mathbb{I}}^{\beta}$. In this case, the mutual information generated in one subunit would have been completely consumed in the other subunit, cf. [**?**, **?**, 16]. Because $\dot{\mathbb{I}}^{channel} \neq 0$, this is, however, no longer valid. Some mutual information for the entire enzyme is additionally generated in the channeling transition involving simultaneously both subunits.

We have computed the values for $\dot{\mathbb{I}}^{\alpha}$, $\dot{\mathbb{I}}^{\beta}$ and $\dot{\mathbb{I}}^{channel}$ under physiological concentrations. We find that they all have the same order of magnitude. The mutual information $\dot{\mathbb{I}}^{channel} = -4.53$ bit s^{-1} generated per unit time by the transition corresponding to indole channeling flows to both subunits where is consumed at the rates of $\dot{\mathbb{I}}^{\alpha} = 3.09$ bit s^{-1} and $\dot{\mathbb{I}}^{\beta} = 1.42$ bit s^{-1}. Note that $\dot{\mathbb{I}}^{\beta}$ is positive whereas all contributions $\iota^{\beta}_{b,b'}$ from individual transitions in the β-subunit are negative. This is an effect of the large contributions from the cross term $\iota^{\beta}_{channel} = 6.43$ bit s^{-1} (whereas $\iota^{\alpha}_{channel} = -0.86$ bit s^{-1}).

Part III

Dynamics of coarse grained spin systems

Chapter 7

Time dependent Renormalization Group transformation and the Kibble Zureck Mechanism

7.1 Critical dynamics and the real space renormalization group

Scaling ideas and the various mathematical formulations of the renormalization group (RG) were developed in the 1960s and early 1970s [32, 60, 9]. They allow a profound understanding of critical singularities near continuous phase transitions in thermal equilibrium. Starting in the mid-1970s, these concepts were subsequently generalized and applied to dynamic critical phenomena.

But even today our knowledge of the dynamic approach towards the equilibrium of a system near a second order phase transition is very limited compared to the achievments in the understanding of static phenomena [32]. In thermal equilibrium the scaling properties of physical quantities, the static critical exponents and amplitudes of the singularities appearing in thermodynamic and correlation functions are accessible. These properties are usually characterised by the dimensionality of the system and the symmetry of its Hamiltonian [26]. The first application of the RG approach in connection with the ϵ-*expansion* to critical dynamics was introduced in 1977 by Halerpin et al. [29, 32]. The dynamics of the system were captured with a Langevin equation of motion by considering a continuous spin field combined with conservations laws. The *real space* RG was first generalized for the study of critical dynamics in 1978 by Achiam and Kosterlitz [1] and Kinzel [34].

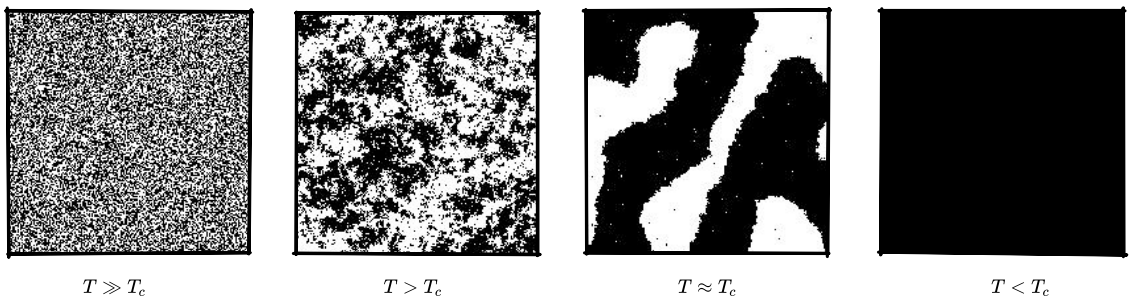

Figure 7.1: Snapshots of a 2D-Ising model adiabatically driven through the second order phase transition from the disordered to the ordered phase.

The application of the RG ideas to calculate time-dependent properties of a system is very similar to their application to static thermal equilibrium. The RG transformation scales the characteristic microscopic length of a systen by a factor b

$$x' = bx. \tag{7.1}$$

This transformation is expressed by a new effective Hamiltonian. A meaningful RG transformation does not alter the physical information in connection with distances longer than the new effective microscopic length. The appearance of a long range correlation length close to a second order phase transition is an example of such a property. The transformation (7.1) can also be applied to the kinetics associated with the relaxation of the system towards thermal equilibrium. Here the RG transformation (7.1) changes the effective microscopic time scale

$$\tau_0' = b^z \tau_0. \tag{7.2}$$

To ensure scale invariance, this has to be compensated for by a contribution from the space sacle to the macroscopic time sacle. Close to the critical point the macroscopic characteristic length is the correlation length ξ. Thus we can relate the scaling of the functional dependence of the macroscopic time scale to ξ. Then we can generally conclude that if the dynamics are under the scaling of both space (7.1) and time (7.2), the characteristic time scale of of the system depends asymptotically on the correlation length, i.e.

$$\tau_c \sim \xi^z, \tag{7.3}$$

where z is the dynamic exponent characterizing the divergence of the typical relaxation time near criticality.

In the analysis of static critical phenomena the RG transformation is usually applied to the equilibrium state probability distribution $p^{eq}(\boldsymbol{\sigma})$. Therefore all results obtained are automatically applicable to all other physical quantities which are related to p^{eq}. The natural generalization for the dynamics is to study the RG transformation of the time-dependent probability distribution $p(\boldsymbol{\sigma}, t)$ and in particular its associated ME (4.2).

7.2 Time dependent renormalization group transformation of a spin system master equation

We consider a system with many degrees of freedom, e.g. a set of $N \gg 1$ spins denoted by $\boldsymbol{\sigma}$, distributed according to the *time dependent* Boltzmann weights

$$p(\boldsymbol{\sigma}, t) = \frac{1}{Z(t)} \exp\left(-\mathcal{H}[\boldsymbol{\sigma}, \mathrm{t}]\right). \tag{7.4}$$

For the effective Hamiltonian we use a simplest possible short-range interaction that involves only neighboring spins.

$$\mathcal{H}[\boldsymbol{\sigma}, t] \coloneqq \mathrm{K}_\beta(t) \sum_{\langle i,j \rangle} \hat{S}(\boldsymbol{\sigma}), \tag{7.5}$$

where $\hat{S}(\boldsymbol{\sigma})$ is the spin operator and $\mathrm{K}_\beta(t)$ the time dependent bath coupling. We are interested in the analysis of a ME that generates configurations that are weighted according to the distribution in Eq. (7.4). Generally this is achived by a ME of an immense state space, i.e. with increasing values of N, the number of spin configurations is growing as q^N if each spin can be in one of q states. This giant non-Markovian master equation has the simple form

$$\tau_0 \frac{\mathrm{d}}{\mathrm{d}t} p(\boldsymbol{\sigma}, t) = \sum_m W_{\boldsymbol{\sigma}^m \boldsymbol{\sigma}}(t) p(\boldsymbol{\sigma}, t) - W_{\boldsymbol{\sigma} \boldsymbol{\sigma}^m}(t) p(\boldsymbol{\sigma}^m, t), \tag{7.6}$$

where $\boldsymbol{\sigma}^m \equiv \hat{f}^m(\boldsymbol{\sigma})$ denotes a configuration where the m-th spin is flipped by the spin flip operator $\hat{f}^m$ and τ_0 sets the bare time scale of the bath coupling.. A central assumption is, that we limit the analysis to single spin flip dynamics but leave the explicit form of the rates being only determined by the condition of local detailed

balance and the time-dependent generalised Hamiltonian (7.5). Then it suffices to define the rates like

$$W_{\boldsymbol{\sigma}^m\boldsymbol{\sigma}}(t) \coloneqq \left(\frac{p(\boldsymbol{\sigma}^m,(t)}{p(\boldsymbol{\sigma},t)}\right)^{\frac{1}{2}} = \exp\left(-\frac{1}{2}(\mathcal{H}[\boldsymbol{\sigma}^m,t]-\mathcal{H}[\boldsymbol{\sigma},t])\right). \tag{7.7}$$

Generally it is a very difficult task to determine the transition rates from first principles such as the microscoplic equation of motion resulting fron the Hamiltonian (c.f. chapter 1). Therefor in many cases empirical master equations are used which, hopefully, describe the physical basis of the dynamics. The most popular empirical rate is the Glauber model, thanks to its simplicity [25, 32]. This model reflects the relaxation of an Ising-like system via an interaction with a heat bath. The Glauber model is equivalent to the model A in the notation of Hohenberg et al. [29, 32].

Renormalization group transformation

A RG transformation (such as decimation or block-spin transformation) on the master equation (7.6) should produce a new equation of a similar form for a renormalized probability distribution. Generally a RG transformation $\hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}]$ maps the original spin configuration $\boldsymbol{\sigma}$ to a new one $\boldsymbol{\mu}$, with space scaled by b. The application of the RG transformation to the probability distribution $p(\boldsymbol{\sigma},t)$ generates a new distribution $p(\boldsymbol{\mu},t)$ of a similar spin system, i,e,

$$p(\boldsymbol{\mu},t) = \sum_{\boldsymbol{\sigma}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}]p(\boldsymbol{\sigma},t) \tag{7.8}$$

where $\hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}]$ has to preserve normality

$$\sum_{\boldsymbol{\mu}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}] = 1, \tag{7.9}$$

has to be non-negative

$$\hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}] \geq 0 \tag{7.10}$$

and it should not change the symmetry of the lattice. The transformed probability distribution can similarly be represented by an Hamiltonian with scaled couplings

$$\mathcal{H}[\boldsymbol{\mu},t] = \mathrm{K}'_\beta(t)\sum_{\langle i,j\rangle}\hat{S}(\boldsymbol{\mu}). \tag{7.11}$$

Thus, the RG transformation $\hat{\mathfrak{T}}_b$ can be considered to be similar to the well known static transformation $\hat{\mathfrak{R}}_b$ of the parameter space $\boldsymbol{K}'_\beta = \hat{\mathfrak{R}}_b \boldsymbol{K}_\beta$ where the fixed point of this transformation is associated with a critical point of a satic system. The RG transformation of time is obtained by applying $\hat{\mathfrak{T}}_b$ to the master equation [1, 34]

$$\sum_{\boldsymbol{\sigma}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}]\Big(\tau_0 \frac{\mathrm{d}}{\mathrm{d}t} p(\boldsymbol{\sigma},t)\Big) = \sum_{\boldsymbol{\sigma}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}]\Big(\sum_m W_{\boldsymbol{\sigma}^m\boldsymbol{\sigma}}(t) p(\boldsymbol{\sigma},t) - W_{\boldsymbol{\sigma}\boldsymbol{\sigma}^m}(t) p(\boldsymbol{\sigma}^m,t)\Big)$$

The RG transformation is time independent and therefore commutes with the time derivative in (7.12), i.e.

$$\sum_{\boldsymbol{\sigma}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}]\Big(\frac{\mathrm{d}}{\mathrm{d}t} p(\boldsymbol{\sigma},t)\Big) = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{\boldsymbol{\sigma}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}] p(\boldsymbol{\sigma},t) = \frac{\mathrm{d}}{\mathrm{d}t} p(\boldsymbol{\mu},t), \tag{7.12}$$

while the transformation of the probability currents is more tedious. Ideally we would expect a result with transformed transition rates and time scale

$$\sum_{\boldsymbol{\sigma}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu},\boldsymbol{\sigma}]\Big(\sum_m W_{\boldsymbol{\sigma}^m\boldsymbol{\sigma}}(t) p(\boldsymbol{\sigma},t) - W_{\boldsymbol{\sigma}\boldsymbol{\sigma}^m}(t) p(\boldsymbol{\sigma}^m,t)\Big) \tag{7.13}$$

$$\mapsto \quad b^{-z} \sum_m W_{\boldsymbol{\mu}^m\boldsymbol{\mu}}(t) p(\boldsymbol{\mu},t) - W_{\boldsymbol{\mu}\boldsymbol{\mu}^m}(t) p(\boldsymbol{\mu}^m,t). \tag{7.14}$$

When a time dependent RG transformation (7.14) of the ME can be computed, then the dynamic RG transformation is accompanied by the scaling of the bare bath coupling time scale

$$\tau'_0 = \tau_0 b^z, \tag{7.15}$$

in which z is immediately identified as the dynamic exponent.

7.3 Kibble-Zurek Mechanism

The Kibble-Zurek Mechanism (KZM) considers the dynamics of spontaneous symmetry breaking in the course of a phase transition induced by the change of a control parameter over time, e.g. the temperature $T \coloneqq 1/\beta$ [67, 33, 9]. In the limit where the coupling is close to its critical value $\mathrm{K}_\beta \to \mathrm{K}_{\beta_c}$, the phenomena of critical slowing down prevents the system from exponentially relaxing into the equilibrium state. Within the KZM the freeze out of the system depending on the quench rate is investigated.

According to the scaling hypothesis of static phase transitions, close to the critical point the correlation length ξ is scaled by the dimensionless temperature $\varepsilon \coloneqq |T - T_c|/T_c$ as

$$\xi(\varepsilon) \sim \xi_0 \varepsilon^{-\nu}, \tag{7.16}$$

with a critical exponent ν and the equilibrium relaxation time as

$$\tau_r(\varepsilon) \sim \tau_0 \left(\frac{\xi(\varepsilon)}{\xi_0} \right)^z \sim \tau_0 \varepsilon^{-z\nu}, \tag{7.17}$$

with a critical exponent $\Delta = \nu z$. We assume that system is preparded in the high symmetry phase at $t_0 = -\infty$. At $t > t_0$ the temperature is lowerd according to the linear protocol

$$T(t) \coloneqq T_c \left(1 - \frac{t}{\tau_q} \right), \tag{7.18}$$

with τ_q the quench rate. The reduced temperature accordingly becomes

$$\varepsilon(t) = \frac{|t|}{\tau_q}. \tag{7.19}$$

Before the temperature reaches T_c, the relaxation time becomes longer than the remaining time $|t|$ to the critical point and the dynamics are not adiabatic anymore. The groth of the correlation length stops before the temperature reaches T_c and the system is approximately frozen due to the divergence of the relaxation time. The system is unable to adjust to the externally imposed change of the reduced controll parameter (7.19). The time $\hat{t}$, when the system freezes, is determined by the equality between the relaxation time τ_r and the remaining time to the critical point $|t|$, i.e.

$$\tau_r(\varepsilon(\hat{t})) = |\hat{t}|. \tag{7.20}$$

With eqs. (7.17), (7.19) and $\hat{\varepsilon} \coloneqq \varepsilon(\hat{t})$ we have

$$\tau_0 \hat{\varepsilon}^{-z\nu} \sim \tau_q \hat{\varepsilon}, \tag{7.21}$$

which can equivalently be written as

$$\hat{\varepsilon} \sim \left(\frac{\tau_q}{\tau_0} \right)^{-1/(1+z\nu)}. \tag{7.22}$$

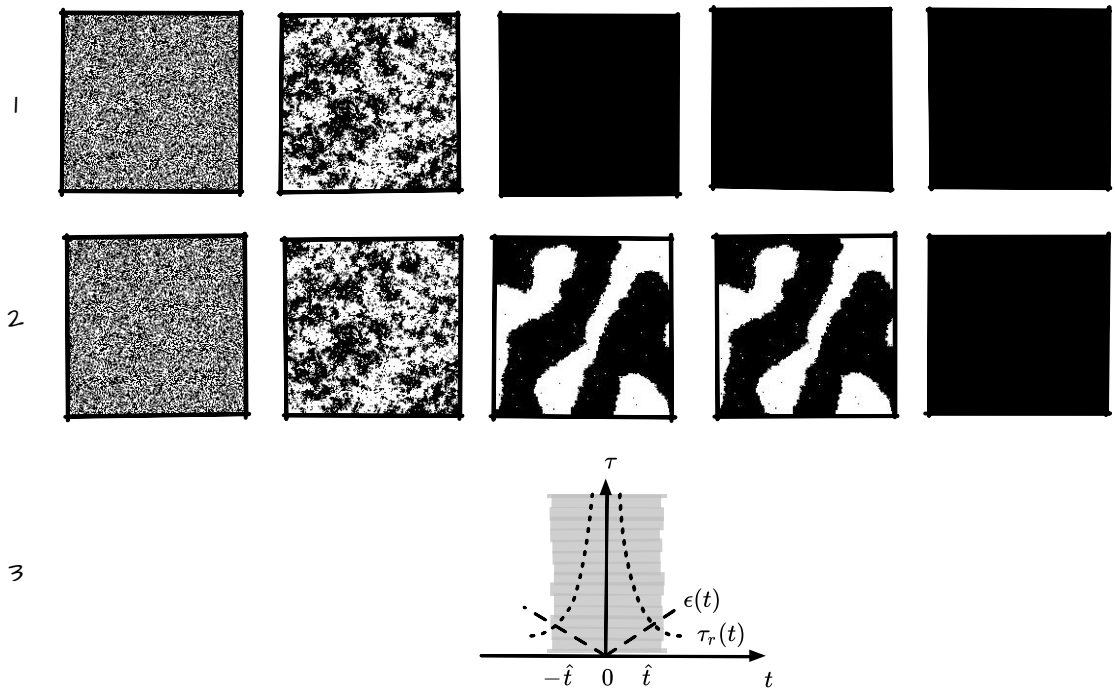

Figure 7.2: **(1)** A system rapidly driven trough the phase transition will quickly fall out of equilibrium once the critical point is reached. **(2)** If the system is driven slower, the adiabatic regime remains for a longer time, but the system will also fall out of equilibrium due to the diverging relaxation time. **(3)** Schematic representation of the linear driving machanism and the timeframe under which the system will be frozen and out of equilibrium.

Finally we can estimate that the correlation length after the quench is scaled by

$$\hat{\xi} \coloneqq \xi(\hat{\varepsilon}) \sim \xi_0 \left(\frac{\tau_q}{\tau_0}\right)^{\nu/(1+z\nu)} . \tag{7.23}$$

At time $|\hat{t}|$ the system freezes and falls out of equilibrium. For the remaining quench time the system is characterized by

$$p^q(\boldsymbol{\sigma}, t) \coloneqq \frac{1}{Z}\exp\left(-\mathcal{H}[\boldsymbol{\sigma}, |\hat{\mathrm{t}}|]\right), \qquad \mathrm{t} \in [-\hat{\mathrm{t}}, \hat{\mathrm{t}}] \tag{7.24}$$

i.e. $\frac{\mathrm{d}}{\mathrm{d}t}p^q(\boldsymbol{\sigma}, t) = 0$ for $t \in [-\hat{t}, \hat{t}]$. At times $|\hat{t}| < t$ the system should not obey detailed balance any more, i.e.

$$W_{\boldsymbol{\sigma}^m\boldsymbol{\sigma}}(t)p^q(\boldsymbol{\sigma}, t) \neq W_{\boldsymbol{\sigma}\boldsymbol{\sigma}^m}(t)p^q(\boldsymbol{\sigma}^m, t), \qquad |\hat{t}| < t \tag{7.25}$$

since the thermodynamic constraints change faster than the system can adapt to it.

Chapter 8

Coarse grained triangular block spin master equation

The last chapter is a summary of preliminary results obtained to develop a systematic RG transformation method combining results from the 80s [29, 1, 34] and current insights from the field of Stochastic Thermodynamics for systems obeying time scale separation [21]. The aim is to approximately incorporate the interactions of the block spins in the transition rate matrix for a low dimensional system desctiption and to obtain a recursion for the transformation of the rate matrix accross scales. The author has the dream to map the critical dynamics of the lattice system onto a single spin under feedback control of two adjacent spins, where the feedback enters into the rates via the perturbative expansion of the block spin intercell interactions in the Boltzmann weights and ideally is allowed to be time dependent for the KZM driving. This might yield an alternative approach to the perturbative treatment of the probability distribution [29, 1, 34] for the computation of the critical exponent z for larger systems and allow new insights into the analysis of the KZM. The author is commited to follow up this idea beyond this thesis.

8.1 Coarse grained transition rates

We focus on the block-spin transformation, where spins $\{\sigma\}$ are divided in to groups $\{\mu\}$ which define new coarse grained spins [32, 9]. A spin belonging to a group μ is

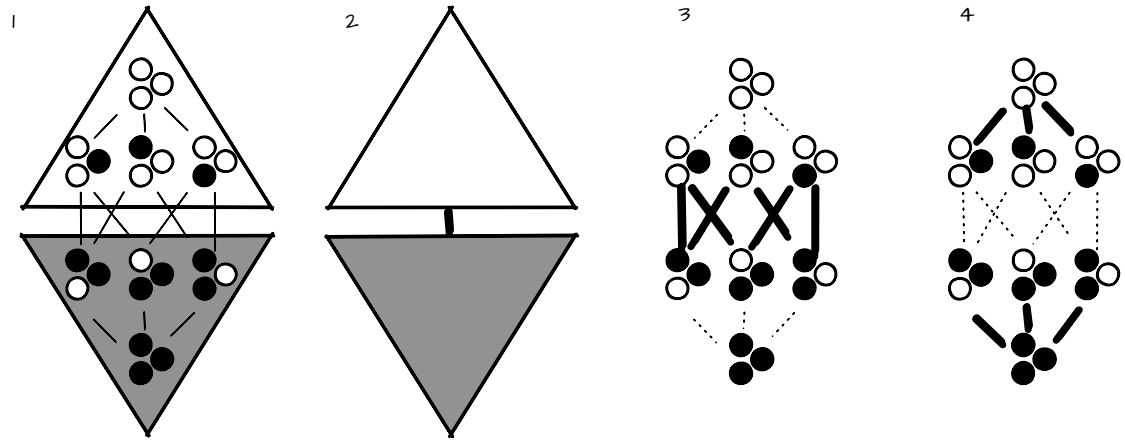

Figure 8.1: **(1)** Only single spin flip dynamics are allowed by assumption. In the example of triangular block spins, the block-spin points up if the majority of micro-spins points up and down, if the majority of micro-spins points down. **(2)** The dynamics of the block-spin are determined by new effective rates. **(3)** Block-spin transitions are realised by spin flips of micro-spins. **(4)** Spin flips not altering the majority of up or down spin states in a block spin, leave the block spin invarian.

denoted by σ_μ. Such a RG transformation defines a new probability distribution

$$p(\boldsymbol{\mu}, t) = \sum_{\boldsymbol{\sigma}} \hat{\mathfrak{T}}_b[\boldsymbol{\mu}, \boldsymbol{\sigma}] p(\boldsymbol{\sigma}, t) \tag{8.1}$$

$$= \sum_{\boldsymbol{\sigma}_\mu} p(\boldsymbol{\sigma}, t). \tag{8.2}$$

Further on we call spins $\sigma \in \boldsymbol{\sigma}$ *micro-spins* and the the coarse grained spins $\mu \in \boldsymbol{\mu}$ *meso-spins* or *block-spins*. Following M. Esposito [21], the time dependent rates for the coarse grained master equartion of (8.2) are given as

$$W_{\boldsymbol{\mu}\boldsymbol{\mu}'}(t) = \sum_{\sigma_{\boldsymbol{\mu}}, \sigma'_{\mu'}} W_{\sigma_{\boldsymbol{\mu}}\sigma'_{\mu'}}(t) p(\boldsymbol{\sigma}'|\boldsymbol{\mu}', t). \tag{8.3}$$

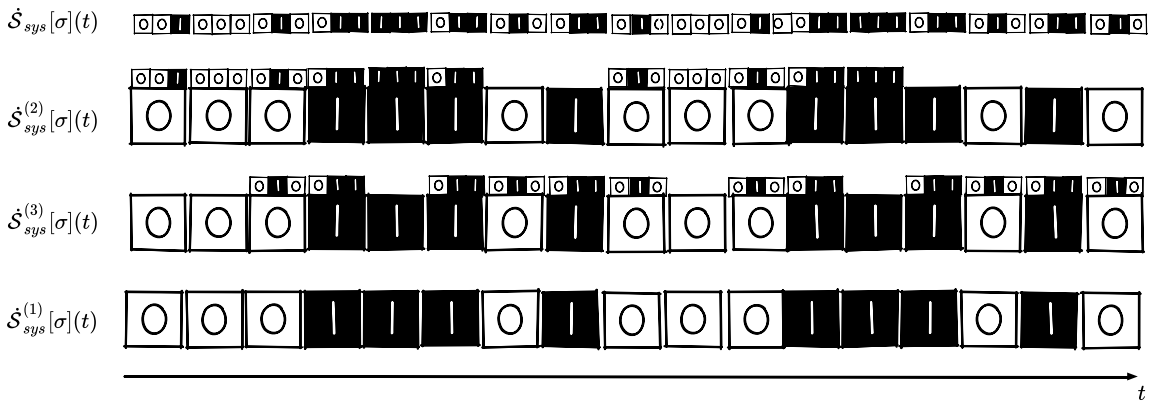

Figure 8.2: Illustration of the block-spin and micro-spin time series reflecting the different entropy production rates.

8.2 Coarse grained entropy production rate

In [21] Esposito seperates the time evolution of the entropy into three contributions, $\dot{\mathbb{H}}[\boldsymbol{\sigma}] = \dot{\mathbb{H}}^{(1)} + \dot{\mathbb{H}}^{(2)} + \dot{\mathbb{H}}^{(3)}$. The part $\dot{\mathbb{H}}^{(2)}$ arises from transitions between configurations of micro-spins belonging ($\boldsymbol{\sigma_\mu} \to \boldsymbol{\sigma'_\mu}$) within a group μ (see figure 8.2 (2)), $\dot{\mathbb{H}}^{(3)}$ due to micro-spin transitions ($\boldsymbol{\sigma_\mu} \to \boldsymbol{\sigma'_{\mu'}}$) connecting two different block-spin configurations (see figure 8.2 (3)) and $\dot{\mathbb{H}}^{(1)}$ due to the effective dynamics of block-spin configurations ($\boldsymbol{\mu} \to \boldsymbol{\mu'}$) (see figure 8.2 (4)). The entropy production follows the same seperation (see Appendix D.1 for derivation)

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}](t) = \dot{\mathcal{S}}^{(1)}_{sys}[\boldsymbol{\sigma}](t) + \dot{\mathcal{S}}^{(2)}_{sys}[\boldsymbol{\sigma}](t) + \dot{\mathcal{S}}^{(3)}_{sys}[\boldsymbol{\sigma}](t). \tag{8.4}$$

On the level of the effective block-spin master equation the visible entropy production would be $\dot{\mathcal{S}}^{(1)}_{sys}[\boldsymbol{\sigma}](t)$. Thus a coarse grained master equation underestimates the entropy production by $\dot{\mathcal{S}}^{(2)}_{sys}[\boldsymbol{\sigma}](t)$ and $\dot{\mathcal{S}}^{(3)}_{sys}[\boldsymbol{\sigma}](t)$ [21]. The hidden entropy production due to the micro-state dynamics, hidden in the meso-states, is $\dot{\mathcal{S}}^{(2)}_{sys}[\boldsymbol{\sigma}](t)$. It is the entropy production arising from internal transitions within a block-spin μ. The three contributions read

$$\dot{\mathcal{S}}^{(1)}_{sys}[\boldsymbol{\sigma}](t) \;\coloneqq\; \sum_{\boldsymbol{\mu},\boldsymbol{\mu'}} W_{\boldsymbol{\mu}\boldsymbol{\mu'}}(t) p(\boldsymbol{\mu'},t) \log \frac{W_{\boldsymbol{\mu'}\boldsymbol{\mu}}(t) p(\boldsymbol{\mu},t)}{W_{\boldsymbol{\mu}\boldsymbol{\mu'}}(t) p(\boldsymbol{\mu'},t)} \geq 0 \tag{8.5}$$

$$\dot{\mathcal{S}}^{(2)}_{sys}[\boldsymbol{\sigma}](t) \;\coloneqq\; \sum_{\boldsymbol{\mu}} p(\boldsymbol{\mu},t) \left(\sum_{\boldsymbol{\sigma_\mu},\boldsymbol{\sigma'_\mu}} W_{\boldsymbol{\sigma_\mu}\boldsymbol{\sigma'_\mu}}(t) p(\boldsymbol{\sigma'_\mu}|\boldsymbol{\mu},t) \log \frac{W_{\boldsymbol{\sigma'_\mu}\boldsymbol{\sigma_\mu}}(t) p(\boldsymbol{\sigma_\mu}|\boldsymbol{\mu},t)}{W_{\boldsymbol{\sigma_\mu}\boldsymbol{\sigma'_\mu}}(t) p(\boldsymbol{\sigma'_\mu}|\boldsymbol{\mu},t)} \right) \geq 0$$

$$\dot{\mathcal{S}}^{(3)}_{sys}[\boldsymbol{\sigma}](t) \;\coloneqq\; \sum_{\boldsymbol{\mu},\boldsymbol{\mu'}} W_{\boldsymbol{\mu}\boldsymbol{\mu'}}(t) p(\boldsymbol{\mu'},t) R_{\boldsymbol{\mu}\boldsymbol{\mu'}}(t) \geq 0. \tag{8.6}$$

The relative entropy

$$R_{\boldsymbol{\mu}\boldsymbol{\mu}'}(t) \;\coloneqq\; \sum_{\boldsymbol{\sigma}_{\boldsymbol{\mu}},\boldsymbol{\sigma}'_{\boldsymbol{\mu}'}} p(\boldsymbol{\sigma}_{\boldsymbol{\mu}} \to \boldsymbol{\sigma}'_{\boldsymbol{\mu}'}|\boldsymbol{\mu} \to \boldsymbol{\mu}', t) \log \frac{p(\boldsymbol{\sigma}_{\boldsymbol{\mu}} \to \boldsymbol{\sigma}'_{\boldsymbol{\mu}'}|\boldsymbol{\mu} \to \boldsymbol{\mu}', t)}{p(\boldsymbol{\sigma}'_{\boldsymbol{\mu}'} \to \boldsymbol{\sigma}_{\boldsymbol{\mu}}|\boldsymbol{\mu}' \to \boldsymbol{\mu}, t)}, \quad (8.7)$$

can be interpreted as an entropic force arising from the micro-state dynamical randomness at which transisitons between block-spin-states occur. The conditional probability that if jumps ($\boldsymbol{\mu} \to \boldsymbol{\mu}'$) occur it is due to a micro-state transition ($\boldsymbol{\sigma}_{\boldsymbol{\mu}} \to \boldsymbol{\sigma}'_{\boldsymbol{\mu}'}$) is given by

$$p(\boldsymbol{\sigma}_{\boldsymbol{\mu}} \to \boldsymbol{\sigma}'_{\boldsymbol{\mu}'}|\boldsymbol{\mu} \to \boldsymbol{\mu}', t) \coloneqq \frac{W_{\boldsymbol{\sigma}_{\boldsymbol{\mu}}\boldsymbol{\sigma}'_{\boldsymbol{\mu}'}}(t) p(\boldsymbol{\sigma}', t)}{W_{\boldsymbol{\mu}\boldsymbol{\mu}'}(t) p(\boldsymbol{\mu}', t)}. \quad (8.8)$$

The entropy production rate is scale invariant only if $\dot{\mathcal{S}}^{(2)}_{sys}[\boldsymbol{\sigma}](t)$ and $\dot{\mathcal{S}}^{(3)}_{sys}[\boldsymbol{\sigma}](t)$ vanish. For the KZM, when the system freezes out, it can be conjectured that this is the case if the spin transitions inside a blockspin are described by an equilibrium process obeying detailed balance and therefore $\dot{\mathcal{S}}^{(2)}_{sys}[\boldsymbol{\sigma}](t) = 0$, but also a high long range correlation between block-spin and micro-spin transitions resulting in a vanishing of $R_{\boldsymbol{\mu}\boldsymbol{\mu}'}(t)$.

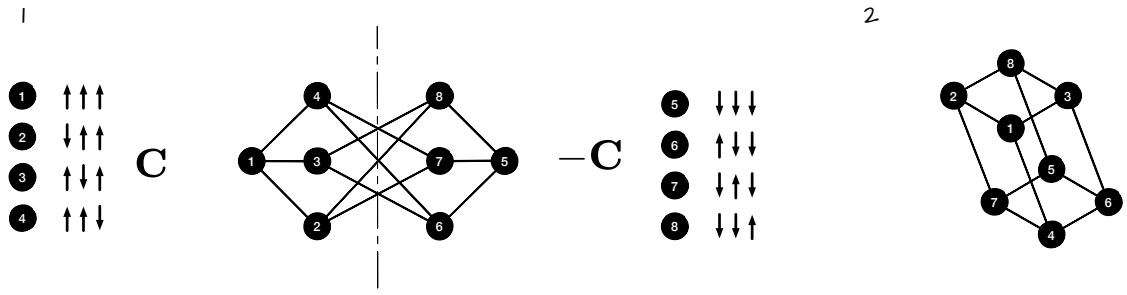

Figure 8.3: **(1)** On the left are configurations which by the majority rule give an ↑ configuration of the blockspin. The components are labeled by the indices $\{1, 2, 3, 4\}$. On the right, after flipping the configuration $\boldsymbol{C}$, are all components $\{5, 6, 7, 8\}$ making up the ↓ configuration. The graph displays the adjacency matrix of the ME and the micro-spin transitions traversing the dotted line, i.e. connecting $\boldsymbol{C}$ with $-\boldsymbol{C}$ and flipping the block-spin. **(2)** The ME also obeys a bipartite structure thanks to the single spin flip dynamics.

8.3 RG transformation of a block spin ME onto a single spin ME

As already noted above, the ME rapidly grows with the number of involves spins, since the number of possible spin configurations and therefore the state space grows as q^N, where q is the number of states the spin can be in and N the number of lattice points. This renders the computation of the steady state probability distribution to be a very complicated task.

All possible configurations $\boldsymbol{C}$ of the blockspin which have an inversion complement $-\boldsymbol{C}$, i.e. the configurations obtained by flipping the hole lattice are given by

$$\boldsymbol{c} \coloneqq \begin{pmatrix} \boldsymbol{C} \\ -\boldsymbol{C} \end{pmatrix}, \qquad \boldsymbol{C} \coloneqq \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}. \tag{8.9}$$

Our aim is to formulate the ME in terms of indices representing the configuration of the lattice. As a next step we need to establish all possible transitions in the ME which connect rows of $\boldsymbol{c}$. Single spin flip dynamics are assumed, meaning that only transitions between configurations involving a single spin flip are allowed. We define an adjacency matrix $\boldsymbol{M}^{(2)}$ which holds all micro spin transitions of the ME which

don't change the state of the block-spin, i.e.

$$\boldsymbol{M}^{(2)} \coloneqq \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \boldsymbol{M}^{(2)\intercal}, \tag{8.10}$$

and an adjacency matrix $\boldsymbol{M}^{(3)}$ which holds all micro spin transitions of the ME which flip the state of the block-spin, i.e.

$$\boldsymbol{M}^{(3)} \coloneqq \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} = \boldsymbol{M}^{(2)\intercal}, \tag{8.11}$$

transitions. Therefore the full adjacency matrix $\boldsymbol{M}$ for the entire graph of the ME of spin-configurations reads

$$\boldsymbol{M} \coloneqq \begin{pmatrix} \boldsymbol{M}^{(2)} & \boldsymbol{M}^{(3)} \\ \boldsymbol{M}^{(3)} & \boldsymbol{M}^{(2)} \end{pmatrix} = \boldsymbol{M}^{\intercal}. \tag{8.12}$$

The RG is implemented with a vector

$$\boldsymbol{t} \coloneqq (\begin{pmatrix} \mathbf{1} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{1} \end{pmatrix})^{\intercal}, \qquad \mathbf{1} \coloneqq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{8.13}$$

which transforms the micro-spin configuration probability vector $\boldsymbol{p}$ into the blockspin probability vector $p_\alpha^{(1)} = \boldsymbol{t}_\alpha \boldsymbol{p}$, and the ME accordingly

$$\frac{\mathrm{d}}{\mathrm{d}t} p_\alpha^{(1)} = \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{t}_\alpha \boldsymbol{p} = \boldsymbol{t}_\alpha \boldsymbol{M} \boldsymbol{W} \boldsymbol{p} = \sum_{\alpha'} \sum_{i_\alpha j_{\alpha'}} W_{i_\alpha j_{\alpha'}} p_{j_{\alpha'}}. \tag{8.14}$$

We note that here in this simple case:

$$\sum_{i_\alpha j_\alpha} W_{i_\alpha j_{\alpha'}} p_{j_{\alpha'}} = \boldsymbol{M}^{(2)} \boldsymbol{W} \boldsymbol{p} \tag{8.15}$$

$$\sum_{\alpha' \neq \alpha} \sum_{i_\alpha j_{\alpha'}} W_{i_\alpha j_{\alpha'}} p_{j_{\alpha'}} = \boldsymbol{M}^{(3)} \boldsymbol{W} \boldsymbol{p} \tag{8.16}$$

8.3.1 Coarse grained transition rates for the sinle spin

In order to determine the transition rates and equilibrium probability distribution we start by introducting a coupling matrix for the computation of the nearest neighbor coupling in the micro-spin Hamiltonian

$$\boldsymbol{K}_0 \coloneqq \frac{\mathrm{K}}{2}\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix} = \boldsymbol{K}_0^{\intercal}. \tag{8.17}$$

With this definition we can compute a matrix, which on its diagonal holds the Hamiltonians of all intracell configurations of $\boldsymbol{c}$, i.e.

$$\boldsymbol{H} \coloneqq \mathrm{diag}(\boldsymbol{c}\boldsymbol{K}_0\boldsymbol{c}^{\intercal}). \tag{8.18}$$

The Hamiltonian of a particular configuration i is then given by the matrix element $[\boldsymbol{H}]_{ii}$. With the use of these matrix elements we can determine the probability to be in a configuration represented by row i of $\boldsymbol{c}$, namely

$$p_i = \frac{1}{Z}\left[e^{-\boldsymbol{H}}\right]_{ii}, \qquad Z \coloneqq \mathrm{Tr}\left\{e^{-\boldsymbol{H}}\right\}. \tag{8.19}$$

Now it is straight forward to compute the transition rates defined by the local detailed balance conditions, i.e.

$$W_{ij} \;\coloneqq\; \left(\frac{p_j}{p_i}\right)^{\frac{1}{2}} = \left[e^{-\frac{1}{2}\boldsymbol{H}}\right]_{jj}\left[e^{\frac{1}{2}\boldsymbol{H}}\right]_{ii}. \tag{8.20}$$

Since (8.17) is symmetric the rates are also symmetric. The oarse grained transition rates for a single blockspins are

$$W^{(1)}_{\alpha'\alpha} \;=\; \sum_{ij}\left[\boldsymbol{t}_\alpha\right]_{ij} M_{ji} W_{ij}\frac{p_j}{p^{(1)}_n} \tag{8.21}$$

where $p^{(1)}_\alpha \coloneqq \sum_i \left[\boldsymbol{t}_\alpha\right]_{ii} p_i$. As introduced in section 8.1 with equation (8.3) the blockspin transition rates are the sum of all possible micro-spin flips resulting in a blockspin flip, but weighted with the conditional probability $p_i/p^{(1)}_\alpha$.

8.4 Feedback through intercell coupling

The rates connecting the up and down state of a single block spin are symmetric since the flip of a spin doesn't change the energy. This symmetry is broken once the

block-spin is coupled to another blockspin. The coupling changes the energy barrier of some spin flips inside a blockspin. This can be interpreted as a mutual feedback between single spins and the nearest neighbor coupling through the entire spin-lattice. The idea is to write the feedback rates as

$$W^{\beta\gamma}_{\alpha'\alpha} \coloneqq W^{(1)}_{\alpha'\alpha} \left(\frac{\exp\left(U^{\beta\gamma}_{\alpha}\right)}{\exp\left(U^{\beta\gamma}_{\alpha'}\right)} \right)^{\frac{1}{2}} \qquad U^{\beta\gamma}_{\alpha} \coloneqq K^{(1)} \left(\sigma_\alpha \sigma_\beta + \sigma_\alpha \sigma_\gamma\right). \tag{8.22}$$

where $K^{(1)} = \frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}$ is the well known renormalized coupling constant [32] and β, γ label the adjacent block spins to α. Ideally it would be possible to shown that this form of the rates can be obtained by a perturbative expansion of the interactions in the rates, in the same way as it is applied to compute the renormalized Hamiltonian in the static RG transformation.

Part IV

Summary

Information thermodynamics

In this study, methods of stochastic thermodynamics have been applied to characterize the operation of the channeling enzyme tryptophan synthase.

Using thermodynamic identities related to the detailed balance, the Gibbs energy landscape of this enzyme along its main catalytic pathway could be reconstructed from the experimental data. We found that, under *in vivo* conditions, the cycle of this enzyme is driven by the Gibbs energy gradient of approximately 19.56 k_BT between its substrates and products. Thus, under physiological substrate and product concentrations, the enzyme operation is far from thermal equilibrium.

Inside the cycle of tryptophan synthase, only the first substrate binding transitions are thermally activated, with activation energies about 1 k_BT. All other transitions, including the events of product release, correspond to a decrease in the Gibbs energy. In particular, channeling is driven by the energy difference of 5.4 k_BT and does therefore not represent a diffusion process.

Because the enzyme operates far from equilibrium, entropy is persistently produced. We found that 27.79 bits of entropy are produced and the same amount of entropy is exported, on the average, to the environment within one catalytic cycle. The distribution of entropy production over the Markov network is largely nonuniform.

Information interactions between the two catalytic subunits of the enzyme have been analyzed. Both the allosteric interactions between the subunits and the channelling of an intermediate product from one of them to another contribute to the change of mutual information. Thus, the previously existing theory [30] had to be generalized to the situations where, in addition to regulatory interactions between the subsystems, the transitions simultaneously changing the states of both of them can also take place. We found that mutual information is generated both in α- and β-subunits at the rates 3.09 and 1.49 bits per second. This mutual information is consumed in the channeling transition so that the balance is maintained. Moreover, contributions from individual allosterically regulated transitions in each of the subunits to the total mutual information change were determined too.

Thus, we have demonstrated that, through the use of stochastic thermodynamics, a rich quantitative characterization of the nonequilibrium operation of an enzyme can be produced. It would be interesting to perform analogous investigations for other enzymes with several catalytic subunits. Such further investigations can clarify the

connections between various thermodynamic properties of such nanomachines and the aspects of the chemical function of the enzymes.

Relaxation of coarse grained spin systems

A method for the RG transformation of the ME rates of a single, non-interacting, blockspin has been introduced. This methods allows the straight forward recursive RG transformation of a blockspin. Taking the interactions with adjacent blockspins into account is a more difficult task, especially if the state-space remain the same. A possible approach is to try a perturbative expansion of the interaction, similar to the one in the static RG transformation. Once the feedback-rates are obtain, the full machinery of Information Thermodynamics can be applied and the influence of driving (KZM) can be investigated. The author is commited to follow up on this direction of studies.

Part V

Appendix

Appendix A

Conversation with a Slow Student

"Before closing this chapter, I want to come back to the thing that troubled Einstein so deeply. I don't know for sure, but I suspect that it had to do with the ultimate meaningless nature of probabilistic statements. I have always been mystified by what they actually say about the world. As far as I can tell, they don't say anything very definite. I once wrote the following very short story, originally included in John Brockman's book What We Believe but Cannot Prove, that illustrates the point. The story, "Conversation with a Slow Student" is about a discussion between a physics professor and a student who just can't get the point. When I wrote the story, I was thinking of myself as the student, not the professor.

Student: Hi Prof. I've got a problem. I decided to do a little probability experiment—you know, coin flipping—and check some of the stuff you taught us. But it didn't work.

Professor: Well I'm glad to hear that you're interested. What did you do?

Student: I flipped this coin 1,000 times. You remember, you taught us that the probability to flip heads is one half. I figured that meant that if I flip 1,000 times I ought to get 500 heads. But it didn't work. I got 513. What's wrong?

Professor: Yeah, but you forgot about the margin of error. If you flip a certain number of times then the margin of error is about the square root of the number of flips. For 1,000 flips the margin of error is about 30. So you were within the margin of error.

Student: Ah, now I get it. Every time I flip 1,000 times I will always get something between 470 and 530 heads. Every single time! Wow, now that's a fact I can count on.

Professor: No, no! What it means is that you will probably get between 470 and 530.

Student: You mean I could get 200 heads? Or 850 heads? Or even all heads?

Professor: Probably not.

Student: Maybe the problem is that I didn't make enough flips. Should I go home and try it 1,000,000 times? Will it work better?

Professor: Probably.

Student: Aw come on Prof. Tell me something I can trust. You keep telling me what probably means by giving me more probablies. Tell me what probability means without using the word probably.

Professor: Hmmm. Well how about this: It means I would be surprised if the answer were outside the margin of error.

Student: My god! You mean all that stuff you taught us about statistical mechanics and Quantum Mechanics and mathematical probability: all it means is that you'd personally be surprised if it didn't work?

Professor: Well, uh . . .“

Excerpt from: *Leonard Susskind* The Black Hole War: My Battle With Stephen Hawking to Make the World Safe for Quantum Mechanics.

Appendix B

Combinatorial origin of Shannon's entropy formula

B.1 Entropy of an ensemble of sequences

Stirling's formula is an approximation for the limit $\lim_{N\to\infty} \log N!$ and can be obtained by a saddle point approximation of the integral representation of $N!$ (see *Kardar* p.49 [32]).

$$\lim_{N\to\infty} \log N! = N\log N - N + \mathcal{O}(\log N) \tag{B.1}$$

We use this result to approximate our combinatorical number of all possibilities of realisations $\{\boldsymbol{\sigma}_t\}$ we cannot distinguish from chapter 2.1

$$\begin{aligned}
& \log \mathbb{W}(\{\boldsymbol{\sigma}_0^m\}) && \text{(B.2)}\\
= \quad & \log\left(N! \prod_{\{\boldsymbol{\sigma}_0^m\}} \frac{1}{N_{\boldsymbol{\sigma}_t}!}\right) && \text{(B.3)}\\
= \quad & \log N! - \sum_{\{\boldsymbol{\sigma}_t\}} \log N_{\boldsymbol{\sigma}_t}! && \text{(B.4)}\\
\underset{N\to\infty}{=} \quad & \log N! - \sum_{\{\boldsymbol{\sigma}_t\}} \log\Big(p(\boldsymbol{\sigma}_t)N\Big)! && \text{(B.5)}\\
\sim \quad & N\log N - N - \sum_{\{\boldsymbol{\sigma}_t\}} \Big(p(\boldsymbol{\sigma}_t)N \log\Big(p(\boldsymbol{\sigma}_t)N\Big) - p(\boldsymbol{\sigma}_t)N\Big) && \text{(B.6)}\\
= \quad & N\log N - N - N\sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t)\Big(\log p(\boldsymbol{\sigma}_t) + \log N\Big) + \sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t)N && \text{(B.7)}\\
= \quad & -N\sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t)\log p(\boldsymbol{\sigma}_t) && \text{(B.8)}
\end{aligned}$$

B.2 Units of information

We can change the units by changing the base

$$\log_x \xi = \log_x [y^{\log_y \xi}] = \log_x y \log_y \xi. \tag{B.9}$$

Thus different bases for the logarithm result in information measures which are just constant multiples of each other. Using $\log_2$ units we measure in *bits* and $\log \equiv \log_e$ are units called *nats*.

B.3 Kullback-Leibler Divergence

From (2.11) and with the use of (B.1) we obtain

$$\begin{aligned}
& \log \mathbb{P}(\{\boldsymbol{\sigma}_t\}) && \text{(B.10)}\\
= \quad & \log \Big[N! \prod_{\{\boldsymbol{\sigma}_t\}} \frac{p(\boldsymbol{\sigma}_t^R)^{N_{\boldsymbol{\sigma}_t}}}{N_{\boldsymbol{\sigma}_t}!} \Big] && \text{(B.11)}\\
= \quad & \log \Big[\mathbb{W}(\boldsymbol{\sigma}_t, N) \prod_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t^R)^{N_{\boldsymbol{\sigma}_t}} \Big] && \text{(B.12)}\\
= \quad & \log \mathbb{W}(\boldsymbol{\sigma}_t, N) + \log \prod_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t^R)^{N_{\boldsymbol{\sigma}_t}} && \text{(B.13)}\\
= \quad & \log \mathbb{W}(\boldsymbol{\sigma}_t, N) + \sum_{\{\boldsymbol{\sigma}_t\}} N_{\boldsymbol{\sigma}_t} \log p(\boldsymbol{\sigma}_t^R) && \text{(B.14)}\\
\underset{N\to\infty}{=} \quad & -N \sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t) \log p(\boldsymbol{\sigma}_t) + \sum_{\{\boldsymbol{\sigma}_t\}} N_{\boldsymbol{\sigma}_t} \log p(\boldsymbol{\sigma}_t^R) && \text{(B.15)}
\end{aligned}$$

$$\begin{aligned}
= \quad & -N \sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t) \log p(\boldsymbol{\sigma}_t) + N \sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t) \log p(\boldsymbol{\sigma}_t^R) && \text{(B.16)}\\
= \quad & -N \sum_{\{\boldsymbol{\sigma}_t\}} p(\boldsymbol{\sigma}_t) \log \frac{p(\boldsymbol{\sigma}_t)}{p(\boldsymbol{\sigma}_t^R)} && \text{(B.17)}
\end{aligned}$$

Appendix C

Properties of time discrete Markov chains

C.1 Master equation as a linear map

For an discrete time update from time t to time $t + \Delta t$ the state σ of the system is mapped to a new state σ' with a transition probability $T_{\sigma\sigma'}$ which obeys

$$\sum_{\sigma} T_{\sigma'\sigma} = 1 \tag{C.1}$$

for probability conservation. The time evolution of the states is thus governed by

$$p(\sigma)_{t+\Delta t} = \sum_{\sigma'} T_{\sigma\sigma'} p_t(\sigma'), \tag{C.2}$$

which can, with the use of (C.1), be equally written as

$$p(\sigma)_{t+\Delta t} - p_t(\sigma) = \sum_{\sigma'} T_{\sigma\sigma'} p_t(\sigma') - p_t(\sigma) \sum_{\sigma} T_{\sigma'\sigma}. \tag{C.3}$$

The evolution equation can be writen in a more compact form as a linear map

$$\boldsymbol{p}_{t+\Delta t} = \boldsymbol{T}\boldsymbol{p}_t \tag{C.4}$$

where $\boldsymbol{p}_t$ denotes the vector whose entries are the probabilities $p_t(x)$. The so-called Transfer matrix iterates $\boldsymbol{p}_t$ forward in time and has matrix elements given by

$$\Big[\boldsymbol{T}\Big]_{\sigma\sigma'} \equiv \begin{cases} 1 - \sum_{\sigma \neq \sigma'} T_{\sigma\sigma'}, & x = x' \\ T_{\sigma\sigma'}, & x \neq x' \end{cases} \tag{C.5}$$

C.2 Convergence towards the steady state

With $\boldsymbol{V} \equiv (\boldsymbol{v}_0, \boldsymbol{v}_1, ...)$, a matrix of the the eigenvectors, and $\boldsymbol{D} \equiv diag(d_0, d_1, ...)$, a diagonal matrix of the eigenvalues $1 = |d_0| > |d_1| \geq |d_2| \geq ...$, the transfer matrix can be decomposed as

$$\boldsymbol{T} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{-1} \tag{C.6}$$

According to Frobenius Theorem there is a unique steady state vector $\overline{\boldsymbol{p}}$ which coincides with $\boldsymbol{v}_0$. Any initial probability vector can be decomposed as $\boldsymbol{p}_0 = \sum_i c_i \boldsymbol{v}_i$ and a subsequent iteration with the transfer matrix gives

$$\begin{aligned}
\boldsymbol{p}_{m\Delta t} &= \boldsymbol{T}^m \boldsymbol{p}_0 && \text{(C.7)}\\
&= \left[\boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{-1}\right]^m \boldsymbol{p}_0 && \text{(C.8)}\\
&= \boldsymbol{V} d^m \boldsymbol{V}^{-1} \boldsymbol{p}_0 && \text{(C.9)}\\
&= \sum_i d_i^m c_i \boldsymbol{v}_i && \text{(C.10)}\\
&\sim d_0^m \left[c_0 \boldsymbol{v}_0 + c_1 (\frac{d_1}{d_0})^m \boldsymbol{v}_1 + ...\right] && \text{(C.11)}
\end{aligned}$$

Since $\overline{\boldsymbol{p}} = \boldsymbol{v}_0$, the limit $\lim_{m\to\infty} \boldsymbol{p}_{m\Delta t}$ approaches the steady state as m goes to infinity with a speed of the order of d_1/d_0 exponentially. Since $d_0 > d_i$ for all $i > 0$, the relaxation time τ_r is effectively determined by the second largest eigenvalue d_1:

$$\tau_r \sim -(\log d_1)^{-1} \tag{C.12}$$

In practice, diagonalizing $\boldsymbol{T}$ is generally a very challenging task, especially for systems with a big state space.

C.2.1 Entropy production and entropy flow

The entropy of a time discrete Markovian stochastic process is readily given by

$$\mathbb{H}[\boldsymbol{\sigma}]_t \coloneqq -\sum_\sigma p_\sigma(t) \log p_\sigma(t). \tag{C.13}$$

The time variation of (C.13) can be split as [22]

$$\mathbb{H}[\boldsymbol{\sigma}]_{t+\Delta t} - \mathbb{H}[\boldsymbol{\sigma}]_t = \Delta_i \mathcal{S}[\boldsymbol{\sigma}] + \Delta_e \mathcal{S}[\boldsymbol{\sigma}]_t \tag{C.14}$$

with the *entropy production*

$$\Delta_i \mathcal{S}[\boldsymbol{\sigma}]_t \;\coloneqq\; \frac{1}{2}\sum_{\sigma\sigma'}[T_{\sigma'\sigma}p_\sigma(t) - T_{\sigma\sigma'}p_{\sigma'}(t)]\log\frac{T_{\sigma'\sigma}p_\sigma(t)}{T_{\sigma\sigma'}p_{\sigma'}(t)} \geq 0 \tag{C.15}$$

and the *entropy flow*

$$\Delta_e \mathcal{S}[\boldsymbol{\sigma}]_t \;\coloneqq\; -\sum_{\sigma\sigma'} T_{\sigma'\sigma}p_\sigma(t)\log\frac{T_{\sigma'\sigma}p_\sigma(t+\Delta t)}{T_{\sigma\sigma'}p_\sigma(t)} \tag{C.16}$$

since

$$\mathbb{H}[\boldsymbol{\sigma}]_{t+\Delta t} - \mathbb{H}[\boldsymbol{\sigma}]_t \tag{C.17}$$

$$= -\sum_x p_{t+\Delta t}(x)\log p_{t+\Delta t}(x) + \sum_x p_t(x)\log p_t(x) \tag{C.18}$$

$$\underset{(C.1,C.2)}{=} -\sum_{xx'} T_{\sigma\sigma'}p_t(\sigma')\log p_{t+\Delta t}(x) + \sum_{\sigma\sigma'} T_{\sigma'\sigma}p_t(\sigma)\log p_t(x) \qquad \textit{with the use of ME}$$

$$+\underbrace{\sum_{xx'} T_{\sigma\sigma'}p_t(\sigma')\log p_t(x) - \sum_{xx'} T_{\sigma\sigma'}p_t(\sigma')\log p_t(x)}_{=0} \tag{C.19}$$

$$= -\sum_{\sigma\sigma'} T_{\sigma\sigma'}p_t(\sigma')\log\frac{p_{t+\Delta t}(\sigma)}{p_t(\sigma)} \tag{C.20}$$

$$+\sum_{\sigma\sigma'}[T_{\sigma'\sigma}p_t(\sigma) - T_{\sigma\sigma'}p_t(\sigma')]\log p_t(x) \tag{C.21}$$

$$= -\sum_{\sigma\sigma'} T_{\sigma\sigma'}p_t(\sigma')\log\frac{p_{t+\Delta t}(\sigma)}{p_t(\sigma)} \tag{C.22}$$

$$+\frac{1}{2}\sum_{\sigma\sigma'}[T_{\sigma'\sigma}p_t(\sigma) - T_{\sigma\sigma'}p_t(\sigma')]\log p_t(x) \tag{C.23}$$

$$+\frac{1}{2}\sum_{\sigma\sigma'}[T_{\sigma\sigma'}p_t(\sigma') - T_{\sigma'\sigma}p_t(\sigma)]\log p_t(x') \tag{C.24}$$

$$= -\sum_{\sigma\sigma'} T_{\sigma\sigma'}p_t(\sigma')\log\frac{p_{t+\Delta t}(\sigma)}{p_t(\sigma)} \tag{C.25}$$

$$+\frac{1}{2}\sum_{\sigma\sigma'}[T_{\sigma'\sigma}p_t(\sigma) - T_{\sigma\sigma'}p_t(\sigma')]\log p_t(x) \tag{C.26}$$

$$-\frac{1}{2}\sum_{\sigma\sigma'}[T_{\sigma'\sigma}p_t(\sigma) - T_{\sigma\sigma'}p_t(\sigma')]\log p_t(x') \tag{C.27}$$

$$= -\sum_{\sigma\sigma'} T_{\sigma\sigma'}p_t(\sigma')\log\frac{p_{t+\Delta t}(\sigma)}{p_t(\sigma)} + \frac{1}{2}\sum_{\sigma\sigma'}[T_{\sigma'\sigma}p_t(\sigma) - T_{\sigma\sigma'}p_t(\sigma')]\log\frac{p_t(\sigma)}{p_t(\sigma')}$$

$$\underbrace{-\sum_{\sigma\sigma'} T_{\sigma\sigma'}p_t(\sigma')\log\frac{T_{\sigma'\sigma}}{T_{\sigma\sigma'}} + \sum_{\sigma\sigma'} T_{\sigma\sigma'}p_t(\sigma')\log\frac{T_{\sigma'\sigma}}{T_{\sigma\sigma'}}}_{=0} \tag{C.28}$$

$$= -\sum_{\sigma\sigma'} T_{\sigma\sigma'}p_t(\sigma')\log\frac{T_{\sigma'\sigma}p_{t+\Delta t}(\sigma)}{T_{\sigma\sigma'}p_t(\sigma)} \tag{C.29}$$

$$\frac{1}{2}\sum_{\sigma\sigma'}[T_{\sigma'\sigma}p_t(\sigma) - T_{\sigma\sigma'}p_t(\sigma')]\log\frac{T_{\sigma'\sigma}p_t(\sigma)}{T_{\sigma\sigma'}p_t(\sigma')} \tag{C.30}$$

$$= \Delta_e\mathcal{S}_t + \Delta_i\mathcal{S}_t \tag{C.31}$$

Their thermodynamic interpretation will be covered in the case of continuous time processes in the next section. Here we limit the discussion to mentioning that $\Delta_i \mathcal{S}[\boldsymbol{\sigma}]_t = 0$ only if detailed balance holds, i.e.

$$T_{\sigma\sigma'} p_{\sigma'}(t) = T_{\sigma'\sigma} p_{\sigma}(t) \qquad \forall \sigma \in \boldsymbol{\sigma} \tag{C.32}$$

C.2.2 Dynamical randomness and entropy production at steady state

As a consequence of the Markov property, the probability of a path $\boldsymbol{\sigma}_m$ at steady state is given by

$$\overline{p}(\boldsymbol{\sigma}_m) = T_{\sigma_m\sigma_{m-1}} ... T_{\sigma_1\sigma_0} p(\sigma_0) \tag{C.33}$$

It is interesting to see, that for a steady state probability vector the entropy production is equal to the forward and reversed entropy rates of the same process

$$\begin{aligned}
\Delta_i \mathcal{S}[\boldsymbol{\sigma}] &= \dot{\mathbb{H}}^R[\boldsymbol{\sigma}] - \dot{\mathbb{H}}[\boldsymbol{\sigma}] && \text{(C.34)}\\
&= -\sum_{\sigma\sigma'} \overline{p}_\sigma (T_{\sigma'\sigma} \log T_{\sigma\sigma'}) + \sum_{\sigma\sigma'} \overline{p}_\sigma (T_{\sigma'\sigma} \log T_{\sigma'\sigma}) && \text{(C.35)}\\
&= \frac{1}{2} \sum_{\sigma\sigma'} (T_{\sigma'\sigma}\overline{p}_\sigma - T_{\sigma\sigma'}\overline{p}_{\sigma'}) \log \frac{T_{\sigma'\sigma}\overline{p}_\sigma}{T_{\sigma\sigma'}\overline{p}_{\sigma'}} \geq 0 && \text{(C.36)}
\end{aligned}$$

which has been shown by Gaspard [22]. The result that

$$\dot{\mathbb{H}}[\boldsymbol{\sigma}] = \sum_\sigma \overline{p}(\sigma) \Big[-\sum_{\sigma'} T_{\sigma'\sigma} \log T_{\sigma'\sigma} \Big] \tag{C.37}$$

is in depth carried out on pages 64–65 in the book of Cover and Thomas [14]. There it is shown that $\dot{\mathbb{H}}[\boldsymbol{\sigma}] = \lim_{m\to\infty} \mathbb{H}[x_m|\boldsymbol{\sigma}_{0:m-1}]$ and that from the Markov property and steady state condition

$$\begin{aligned}
\lim_{m\to\infty} \mathbb{H}[x_m|\boldsymbol{\sigma}_{0:m-1}] &= \mathbb{H}[x_m|x_{m-1}] && \text{(C.38)}\\
&= \sum_\sigma \overline{p}(\sigma) \Big[-\sum_{\sigma'} T_{\sigma'\sigma} \log T_{\sigma'\sigma} \Big]. && \text{(C.39)}
\end{aligned}$$

With this result we can imideately see that

$$\begin{aligned}
\Delta_i \mathcal{S}(\boldsymbol{\sigma}) &= \dot{\mathbb{H}}^R[\boldsymbol{\sigma}] - \dot{\mathbb{H}}[\boldsymbol{\sigma}] && \text{(C.40)}\\
&= -\sum_{\sigma\sigma'} \overline{p}(\sigma) T_{\sigma'\sigma} \log T_{\sigma\sigma'} + \sum_{\sigma\sigma'} \overline{p}(\sigma) T_{\sigma'\sigma} \log T_{\sigma'\sigma} && \text{(C.41)}\\
&= \sum_{\sigma\sigma'} \overline{p}(\sigma) T_{\sigma'\sigma} \log \frac{T_{\sigma'\sigma}}{T_{\sigma\sigma'}} && \text{(C.42)}\\
&= \frac{1}{2} \sum_{\sigma\sigma'} \overline{p}(\sigma) T_{\sigma'\sigma} \log \frac{T_{\sigma'\sigma}}{T_{\sigma\sigma'}} + \frac{1}{2} \sum_{\sigma\sigma'} \overline{p}(\sigma') T_{\sigma\sigma'} \log \frac{T_{\sigma\sigma'}}{T_{\sigma'\sigma}} \\
&= \frac{1}{2} \sum_{\sigma\sigma'} [T_{\sigma'\sigma} \overline{p}(\sigma) - T_{\sigma\sigma'} \overline{p}(\sigma')] \log \frac{T_{\sigma'\sigma}}{T_{\sigma\sigma'}} && \text{(C.43)}\\
&\quad + \underbrace{\frac{1}{2} \sum_{\sigma\sigma'} [T_{\sigma'\sigma} \overline{p}(\sigma) - T_{\sigma\sigma'} \overline{p}(\sigma')] \log \frac{\overline{p}(\sigma)}{\overline{p}(\sigma')}}_{=0} && \text{(C.44)}\\
&= \frac{1}{2} \sum_{\sigma\sigma'} [T_{\sigma'\sigma} \overline{p}(\sigma) - T_{\sigma\sigma'} \overline{p}(\sigma')] \log \frac{T_{\sigma'\sigma} \overline{p}(\sigma)}{T_{\sigma\sigma'} \overline{p}(\sigma')} && \text{(C.45)}
\end{aligned}$$

Appendix D

Coarse grained entropy production rate

D.1 The spliting of the entropy production rate

To limit the notational complexity, we omit the time dependence in the following,

$$\dot{\mathcal{S}}_{sys}[\boldsymbol{\sigma}] \tag{D.1}$$

$$= \sum_{k,k'} \sum_{\sigma_k,\sigma'_{k'}} \Phi_{x_k x'_{k'}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}}{\Phi_{\sigma'_{k'}\sigma_k}} \tag{D.2}$$

$$= \sum_{k,k'\neq k} \sum_{\sigma_k,\sigma'_{k'}} \Phi_{\sigma_k \sigma'_{k'}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}}{\Phi_{\sigma'_{k'}\sigma_k}} + \sum_{k} \sum_{\sigma_k,\sigma'_{k}} \Phi_{\sigma_k \sigma'_{k}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}}{\Phi_{\sigma'_{k'}\sigma_k}} \tag{D.3}$$

$$= \sum_{k,k'\neq k} \sum_{\sigma_k,\sigma'_{k'}} \Phi_{\sigma_k \sigma'_{k'}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}\Phi_{kk'}\Phi_{k'k}}{\Phi_{\sigma'_{k'}\sigma_k}\Phi_{kk'}\Phi_{k'k}} + \sum_{k} \sum_{\sigma_k,\sigma'_{k}} \Phi_{\sigma_k \sigma'_{k}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}}{\Phi_{\sigma'_{k'}\sigma_k}} \tag{D.4}$$

$$= \sum_{k,k'\neq k} \sum_{\sigma_k,\sigma'_{k'}} \Phi_{\sigma_k \sigma_{k'}} \log \frac{\Phi_{kk'}}{\Phi_{kk'}} + \sum_{k,k'\neq k} \sum_{\sigma_k,\sigma'_{k'}} \Phi_{\sigma_k \sigma'_{k'}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}\Phi_{k'k}}{\Phi_{\sigma'_{k'}\sigma_k}\Phi_{k,k'}} + \sum_{k} \sum_{\sigma_k,\sigma'_{k}} \Phi_{\sigma_k \sigma'_{k}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}}{\Phi_{\sigma'_{k'}\sigma_k}}$$

$$= \underbrace{\sum_{k,k'} \Phi_{kk'} \log \frac{\Phi_{kk'}}{\Phi_{kk'}}}_{\dot{\mathcal{S}}^{(1)}_{sys}} + \underbrace{\sum_{k,k'\neq k} \sum_{\sigma_k,\sigma'_{k'}} \Phi_{\sigma_k \sigma'_{k'}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}\Phi_{k'k}}{\Phi_{\sigma'_{k'}\sigma_k}\Phi_{k,k'}}}_{\dot{\mathcal{S}}^{(3)}_{sys}} + \underbrace{\sum_{k} \sum_{\sigma_k,\sigma'_{k}} \Phi_{\sigma_k \sigma'_{k}} \log \frac{\Phi_{\sigma_k \sigma'_{k'}}}{\Phi_{\sigma'_{k'}\sigma_k}}}_{\dot{\mathcal{S}}^{(2)}_{sys}} \tag{D.5}$$

In particular $\dot{\mathcal{S}}_i^{(2)}[\boldsymbol{\sigma}]$ can be writte as

$$\dot{\mathcal{S}}_{sys}^{(2)}[\boldsymbol{\sigma}] = \sum_k \sum_{\sigma_k,\sigma'_k} \Phi_{\sigma_k\sigma'_k} \log \frac{\Phi_{\sigma'_k\sigma_k}}{\Phi_{\sigma_k\sigma'_k}} \tag{D.6}$$

$$= \sum_k p(k) \sum_{\sigma_k,\sigma'_k} W_{\sigma_k\sigma'_k} \frac{p(\sigma'_k)}{p(k)} \log \frac{W_{\sigma'_k x_k} p(x_k) p(k)}{W_{\sigma_k\sigma'_k} p(\sigma'_k) p(k)} \tag{D.7}$$

$$= \sum_k p(k) \left(\sum_{\sigma_k,\sigma'_k} W_{\sigma_k\sigma'_k} p(\sigma'_k|k) \log \frac{W_{x'_k x_k} p(\sigma_k|k)}{W_{\sigma_k\sigma'_k} p(\sigma'_k|k)} \right) \geq 0 \tag{D.8}$$

In oder to better understand the remaining contribution $\dot{\mathcal{S}}_{sys}^{(3)}[\boldsymbol{\sigma}]$ we define a conditional probability for that if jumps $(k \to k')$ occur it is due to a micro-state transition $(x_k \to x'_{k'})$.

$$p((x_k \to x'_{k'})|(k \to k')) := \frac{W_{x_k x'_{k'}} p(x'|k',t)}{W_{kk'}} = \frac{\Phi_{x_k x'_{k'}}}{\Phi_{kk'}} \tag{D.9}$$

With these probabilities at hand we rewrite

$$\dot{\mathcal{S}}_{sys}^{(3)}[\boldsymbol{\sigma}] = \sum_{k,k'\neq k} \sum_{\sigma_k,\sigma'_{k'}} \Phi_{\sigma_k\sigma'_{k'}} \log \frac{\Phi_{\sigma_k\sigma'_{k'}} \Phi_{k'k}}{\Phi_{\sigma'_{k'}\sigma_k} \Phi_{kk'}} \tag{D.10}$$

$$= \sum_{k,k'\neq k} \Phi_{kk'} \sum_{\sigma_k,\sigma'_{k'}} \frac{\Phi_{\sigma_k\sigma'_{k'}}}{\Phi_{kk'}} \log \frac{\Phi_{\sigma_k\sigma'_{k'}}}{\Phi_{kk'}} \frac{\Phi_{k'k}}{\Phi_{\sigma'_{k'}\sigma_k}} = \sum_{k,k'\neq k} \Phi_{kk'} R_{kk'}, \tag{D.11}$$

which is the result from [21]. At this point we have introduced the relative entropy

$$R_{kk'} := \sum_{\sigma_k,\sigma'_{k'}} \frac{\Phi_{\sigma_k\sigma'_{k'}}}{\Phi_{kk'}} \log \frac{\Phi_{\sigma_k\sigma'_{k'}}}{\Phi_{kk'}} \frac{\Phi_{k'k}}{\Phi_{\sigma'_{k'}\sigma_k}} \tag{D.12}$$

$$= \sum_{\sigma_k,\sigma'_{k'}} p((\sigma_k \to \sigma'_{k'})|(k \to k')) \log \frac{p((\sigma_k \to \sigma'_{k'})|(k \to k'))}{p((\sigma'_{k'} \to \sigma_k)|(k' \to k))} \tag{D.13}$$

Zusammenfassung

asdfgasdfag

Bibliography

[1] Ya'akov Achiam and JM Kosterlitz. Real-space renormalization group for critical dynamics. *Physical Review Letters*, 41(2):128, 1978.

[2] K. S. Anderson, E. W. Miles, and K. A. Johnson. *J. Biol. Chem.*, 266(13):8020, 1991.

[3] KS Anderson, EW Miles, and KA Johnson. Serine modulates substrate channeling in tryptophan synthase. a novel intersubunit triggering mechanism. *Journal of Biological Chemistry*, 266(13):8020–8033, 1991.

[4] F Alexander Bais and J Doyne Farmer. The physics of information. *In Philosophy of Information, 609-684. Eds. Pieter Adriaans and Johan Van Benthem*, 2007.

[5] Andre C Barato, David Hartich, and Udo Seifert. Efficiency of cellular information processing. *New Journal of Physics*, 16(10):103024, 2014.

[6] Thomas RM Barends, Michael F Dunn, and Ilme Schlichting. Tryptophan synthase, an allosteric molecular factory. *Current opinion in chemical biology*, 12(5):593–600, 2008.

[7] Bryson D Bennett, Elizabeth H Kimball, Melissa Gao, Robin Osterhout, Stephen J Van Dien, and Joshua D Rabinowitz. Absolute metabolite concentrations and implied enzyme active site occupancy in escherichia coli. *Nature chemical biology*, 5(8):593–599, 2009.

[8] Patrick Billingsley. Ergodic theory and information. 1965.

[9] Tobias Brandes. Statistische mechanik. *Vorlesungsmanuskript am Institut für Theoretische Physik der TU Berlin, WS14/15*, 2014, 2014.

[10] Paul C Bressloff. *Stochastic processes in cell biology*, volume 41. Springer, 2014.

[11] PS Brzović, Yoshihiro Sawa, C Craig Hyde, Edith W Miles, and Michael F Dunn. Evidence that mutations in a loop region of the alpha-subunit inhibit the transition from an open to a closed conformation in the tryptophan synthase bienzyme complex. *Journal of Biological Chemistry*, 267(18):13028–13038, 1992.

[12] Patrizia Castiglione, Massimo Falcioni, Annick Lesne, and Angelo Vulpiani. *Chaos and coarse graining in statistical mechanics*. Cambridge University Press Cambridge, 2008.

[13] Isaac P Cornfeld, Sergej V Fomin, and Yakov Grigorevic Sinai. *Ergodic theory*, volume 245. Springer Science & Business Media, 2012.

[14] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[15] James P Crutchfield, Christopher J Ellison, Ryan G James, and John R Mahoney. Synchronization and control in intrinsic and designed computation: an information-theoretic analysis of competing models of stochastic computation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(3):037105, 2010.

[16] Giovanni Diana and Massimiliano Esposito. *J. Stat. Mech.*, 2014(4):P04010, 2014.

[17] Michael F Dunn. Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Archives of biochemistry and biophysics*, 519(2):154–166, 2012.

[18] Michael F Dunn, Valentin Aguilar, Peter Brzovic, William F Drewe Jr, Karl F Houben, Catherine A Leja, and Melinda Roy. The tryptophan synthase bienzyme complex transfers indole between the. alpha.-and. beta.-sites via a 25-30. ang. long tunnel. *Biochemistry*, 29(37):8598–8607, 1990.

[19] Michael F Dunn, Dimitri Niks, Huu Ngo, Thomas RM Barends, and Ilme Schlichting. Tryptophan synthase: the workings of a channeling nanomachine. *Trends in biochemical sciences*, 33(6):254–264, 2008.

[20] J-P Eckmann and David Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of modern physics*, 57(3):617, 1985.

[21] Massimiliano Esposito. Stochastic thermodynamics under coarse graining. *Phys. Rev. E 85, 041125*, 2012.

[22] Pierre Gaspard. Time-reversed dynamical entropy and irreversibility in markovian random processes. *Journal of statistical physics*, 117(3-4):599–615, 2004.

[23] Pierre Gaspard. Hamiltonian dynamics, nanosystems, and nonequilibrium statistical mechanics. *Physica A: Statistical Mechanics and its Applications*, 369(1):201–246, 2006.

[24] Pierre Gaspard and Xiao-Jing Wang. Noise, chaos, and (ε, τ)-entropy per unit time. *Physics Reports*, 235(6):291–343, 1993.

[25] Roy J Glauber. Time-dependent statistics of the ising model. *Journal of mathematical physics*, 4(2):294–307, 1963.

[26] Nigel Goldenfeld. Lectures on phase transitions and the renormalization group. 1992.

[27] David Hartich, Andre C Barato, and Udo Seifert. Sensory capacity: An information theoretical measure of the performance of a sensor. *Physical Review E*, 93(2):022116, 2016.

[28] U Seifert Hartich, Barato. Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a maxwell's demon interpretation. *J. Stat. Mech.*, 2014.

[29] Pierre C Hohenberg and Bertrand I Halperin. Theory of dynamic critical phenomena. *Reviews of Modern Physics*, 49(3):435, 1977.

[30] Esposito Horowitz. Thermodynamics with continuous information flow. *Phys. Rev. X*, 2014.

[31] Edwin T Jaynes. *Probability theory: the logic of science*. Cambridge university press, 2003.

[32] Mehran Kardar. *Statistical physics of particles*. Cambridge University Press, 2007.

[33] Thomas WB Kibble. Topology of cosmic domains and strings. *Journal of Physics A: Mathematical and General*, 9(8):1387, 1976.

[34] W Kinzel. Critical dynamics by real-space renormalisation group. *Zeitschrift für Physik B Condensed Matter*, 29(4):361–362, 1978.

[35] Kasper KIRSCHNER, Robert L WISKOCIL, Martha FOEHN, and Laurel REZEAU. The tryptophan synthase from escherichia coli. *European Journal of Biochemistry*, 60(2):513–523, 1975.

[36] N Kishore, Y B Tewari, D L Akers, R N Goldberg, and E W Miles. *Biophys. Chem.*, 73(3):265, 1998.

[37] Andrei Kolmogoroff. Zur theorie der markoffschen ketten. *Mathematische Annalen*, 112(1):155–160, 1936.

[38] Andrei Nikolaevitch Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. In *Dokl. Akad. Nauk SSSR*, volume 124, pages 754–755, 1959.

[39] A. N. Lane and K. Kirschner. *Biochemistry*, 30(2):479, January 1991.

[40] Andrew N. Lane and Kasper Kirschner. *Eur. J. Biochem.*, 129(3):571, March 2005.

[41] Catherine A Leja, Eilika U Woehl, and Michael F Dunn. Allosteric linkages between. beta.-site covalent transformations and. alpha.-site activation and deactivation in the tryptophan synthase bienzyme complex. *Biochemistry*, 34(19):6552–6561, 1995.

[42] Joseph T Lizier, Mikhail Prokopenko, and Albert Y Zomaya. Local information transfer as a spatiotemporal filter for complex systems. *Physical Review E*, 77(2):026110, 2008.

[43] Dimitri Loutchko, Didier Gonze, and Alexander S. Mikhailov. *J. Phys. Chem. B*, 120(9):2179, March 2016.

[44] Alexander Mikhailov and Alexander Yu Loskutov. *Foundations of synergetics II: chaos and noise*, volume 52. Springer Science & Business Media, 2013.

[45] Huu Ngo, Novelle Kimmich, Rodney Harris, Dimitri Niks, Lars Blumenstein, Victor Kulik, Thomas Reinier Barends, Ilme Schlichting, and Michael F Dunn. Allosteric regulation of substrate channeling in tryptophan synthase: modulation of the l-serine reaction in stage i of the β-reaction by α-site ligands. *Biochemistry*, 46(26):7740–7753, 2007.

[46] Robert K Niven. Combinatorial entropies and statistics. *The European Physical Journal B*, 70(1):49–63, 2009.

[47] Matteo Polettini. Geometric and combinatorial aspects of nonequilibrium statistical mechanics. 2012.

[48] Amilcare Porporato, JR Rigby, and Edoardo Daly. Irreversibility and fluctuation theorem in stationary time series. *Physical review letters*, 98(9):094101, 2007.

[49] Mikhail Prokopenko, Joseph T Lizier, and Don C Price. On thermodynamic interpretation of transfer entropy. *Entropy*, 15(2):524–543, 2013.

[50] Samanta Raboni, Stefano Bettati, and Andrea Mozzarelli. Tryptophan synthase: a mine for enzymologists. *Cellular and molecular life sciences*, 66(14):2391–2403, 2009.

[51] Jürgen Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Review of Modern Physics*, 1976.

[52] Thomas Schreiber. Measuring information transfer. *Physical review letters*, 85(2):461, 2000.

[53] Benjamin Schumacher and Michael Westmoreland. *Quantum processes systems, and information*. Cambridge University Press, 2010.

[54] Ya Ge Sinai. On the concept of entropy of a dynamical system. In *Dokl. Akad. Nauk. SSSR*, volume 124, pages 768–771, 1959.

[55] Philipp Strasberg. Thermodynamics and information processing at the nanoscale. 2016.

[56] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press, 2014.

[57] Leonard Susskind and Art Friedman. *Quantum mechanics: the theoretical minimum*, volume 2. Basic Books, 2015.

[58] Leonard Susskind and George Hrabovsky. *The theoretical minimum: what you need to know to start doing physics*. Basic Books, 2014.

[59] Leonard Susskind and James Lindesay. An introduction to black holes, information and the string theory revolution. *World Scientific, Singapore*, 10(5):10, 2005.

[60] Uwe C Täuber. *Critical dynamics: a field theory approach to equilibrium and non-equilibrium scaling behavior*. Cambridge University Press, 2014.

[61] Tânia Tomé and Mário J De Oliveira. *Stochastic dynamics and irreversibility*. Springer, 2015.

[62] Christian Van den Broeck and Massimiliano Esposito. Ensemble and trajectory thermodynamics: A brief introduction. *Physica A: Statistical Mechanics and its Applications*, 418:6–16, 2015.

[63] Christian Van den Broeck et al. Stochastic thermodynamics: a brief introduction. *Physics of Complex Colloids*, 184:155–193, 2013.

[64] Esposito Van den Broeck. Ensemble and trajectory thermodynamics: A brief introduction. *Physica A*, 2015.

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

[66] Eilika Woehl and Michael F Dunn. Mechanisms of monovalent cation action in enzyme catalysis: the tryptophan synthase α-, β-, and $\alpha\beta$-reactions. *Biochemistry*, 38(22):7131–7141, 1999.

[67] Wojciech H Zurek. Cosmological experiments in superfluid helium? *Nature*, 317(6037):505–508, 1985.