A. I. Alikhanian National Science Laboratory

(Yerevan Physics Institute)

Narek H. Martirosyan

Markov processes in thermodynamics and statistics

Thesis for acquiring the degree of candidate of physical-mathematical sciences

in division 01.04.02 (Theoretical Physics)

Scientific supervisor

Doctor of phys.-math. sciences

Prof. N. Akopov

YEREVAN 2017

Contents

List of Figures

Abbreviations

$1D$ 1 dimensional

$2D$ 2 dimensional

CDF Cumulative distribution function

$CLHEP$ Class Library for High Energy Physics

$ECDF$ Empirical cumulative distribution function

ET Effective temperature

$F-D$ Fluctuation-dissipation

$F-P$ Fokker-Planck

HEP High energy physics

$K-S$ Kolmogorov-Smirnov

LCG Linear congruential generator

MC Monte Carlo

$NESS$ Non-equilibrium steady state

PDF Probability distribution function

$PhoCS$ Photothermal Correlation Spectroscopy

$PRNG$ Pseudo-random number generators

SHE Stochastic heat engine

TEG Thermoelectric generator

$TRNG$ True random number generator

Introduction and Motivation

A variety of physical phenomena can be described within a deterministic approach, where no randomness is involved in their time-evolution, i.e. the same initial conditions always lead to the same final result. Typical examples of these are ordinary differential equations, e.g. the Newton's laws of motion, the Schrödinger's equation *etc.* But there are many situations where the deterministic approach is not applicable. Unlike deterministic situations, in stochastic processes the future behaviour of the system is not uniquely defined by its initial conditions. Rather the system's variables are described by a time-dependent probability distribution. The latter does again obey deterministic (normally differential) equations which refers to a suitable stochastic process, i.e. to a collection of random variables $(X(t),\ t \in T)$, where T is the index set. Note that any deterministic process is a special case of some stochastic process in the sense that the deterministic case events occur with probability one [1].

There is a class of deterministic processes which appear to be stochastic but is not, they are modeled by deterministic algorithms and can be over-sensitive to small changes of initial conditions. An important example of such a system is the pseudo-random number generators (PRNG) [2–4] for generating numbers whose statistical properties are close to the properties of random numbers. A large number of these systems appear in chaos theory, a classic example is the butterfly effect exhibited by nonlinear deterministic systems [5]. More generally, depending on the purpose of modeling, one may choose between deterministic and stochastic approaches to the same phenomenon [1, 6–8].

Stochastic processes have countless applications in numerous areas of science such as physics, chemistry [9], biology [10, 11], neuroscience [12, 13], information theory [14], image

recognition [15–17], cryptography [18], ecology [19] *etc.*

Four main factors define stochastic process: state space, transition probabilities between states, index set and the relationship between random variables at different times. Depending on whether the index set is discrete or continuous, processes are called discrete-time or continuous-time. In this work we will focus on continuous-time, discrete-space stochastic processes with the Markov (or memory-less) property: the probability distribution of future states depends on the current distribution and is independent of the history of states. Such processes are called Markov processes after Andrey Markov [20]. The notion of time can be somewhat generalized: in statistical inference the sample's size of the data plays the role of time, i.e. as we acquire more data points, our probabilities change.

The basic example of a physical system with a continuous-time Markovian behaviour is the Brownian motion, i.e. the random movements of particles in a liquid. Though it was observed by a number of scientists, e.g. by Jan Ingenhousz (1785) who is also famous for the discovery of photosynthesis, the phenomenon is named after Robert Brown (1773-1858) whose discovery was due to systematic studies of pollen grains suspended in water. The theory was advanced by Einstein, Smoluchowski and Langevin. The research into Brownian motion greatly contributed into the development of the theory of stochastic processes.

In March 1905 (also known as the miracle year) after explaining the photoelectric effect creating the quantum theory of light, in April, Einstein explained the Brownian motion via random collisions of the particle by molecules of water. His mathematical description of Brownian motion was important since it was evidence for the existence of atoms [21]. The Einstein's theory was confirmed experimentally by Jean Perrin (1870–1942) [21]. The Einstein's theoretical explanation contains itself the concept of Markov processes and its main equations, those by Chapman-Kolmogorov and Fokker-Planck [9, 22]. Independently from Einstein, Smoluchowski studied the Brownian motion via discrete space and discrete time approach. In his approach the Brownian particle at each discrete time-step jumps to the left or to the right, with equal probability. These type of stochastic processes, where only transitions to neighbouring cites are allowed, are called birth-death processes analogous to random walks. The Smoluchowski's approach to random walks coincides with Einstein's fa-

mous diffusion equation $\frac{\partial P(x,t)}{\partial t} = D\ \frac{\partial^2 P(x,t)}{\partial x^2}$, where D is the diffusion constant and $P(x,t)$ is the probability for Brownian particle to be at point x at time t. Finally, Paul Langevin (1908) obtained the same result for the mean-squared displacement of the Brownian particle (as Einstein) by introducing a random force in the Newton's law. The Langevin equation demonstrates the time-scale separation inherent in the Brownian motion: the particle is much slower than the correlation time of the random force. Hence the latter can be modeled via a white noise.

The Markov process is the most commonly used tool for modeling random evolution. Most systems found in nature satisfy the memory-less property, provided that their state-space is suitably enlarged [23]. Markov processes have a huge variety of applications in physics, statistics, chemistry, biology, Internet, genetics *etc*; e.g. an Internet application of Markov chains (discrete-time Markov process) is the PageRank algorithm [24] used by Google to rank pages after searching websites.

When transition rates between states are known then the time-evolution of probabilities of states of a Markovian system are governed by the so-called master equation; a system of first-order (in time) differential equations [9]

$$\frac{dp_i}{dt} = \sum\nolimits_j [\rho_{i\leftarrow j} p_j(t) - \rho_{j\leftarrow i} p_i(t)], \tag{1}$$

where $p_i(t)$ is the probability of a system to be in state i at time t, and $\rho_{i\leftarrow j}$ is the transition rate from j to i. Note that the transition rates may depend on time and must be positive. The transition rates are said to satisfy detailed balance condition (DBC) when for the stationary state p_i^s the equality $\rho_{i\leftarrow j} p_j^s = \rho_{j\leftarrow i} p_i^s$ holds for every pair (i,j), i.e. all probability currents between states vanish in the stationary state.

In physics, Markovian systems are widespread in non-equilibrium thermodynamics and statistical mechanics, where the macroscopic behaviour of a many-particle system is studied through probabilistic description of states. A system in contact with a thermal bath eventually settles into an equilibrium state which is characterized by zero energy flux between the bath and the system. The virtue of thermal equilibrium is that the probability of

finding a system in any micro-state, has a universal form that depends on its energy via the Boltzmann's factor, $e^{-E_n/kT}/Z$, where the partition function $Z = \sum_n e^{-E_n/kT}$ ensures normalization. Z encodes statistical properties of an equilibrium system: thermodynamic variables, such as the total energy, free energy, entropy, pressure *etc*, can be deduced from Z or its derivatives. Much of the equilibrium statistical mechanics is developed and understood by the Boltzmann's distribution, which also forms the groundwork of molecular biology [25, 26]. Though thermodynamics provides an efficient description of equilibrium states, almost all real systems and processes are far from equilibrium, because they do exchange matter, energy, and information with their environment. Thus the main aspect of non-equilibrium systems is their openness. A daily example of this is the Earth, where there is a continuous energy flux from the Sun and re-radiation back into the space. In essence, all biological systems are open and survive via balanced input and output of energy, matter and information [27, 28]. Driven complex fluids, turbulence, fractures, glasses, foams, colloidal suspensions *etc* do demonstrate non-equilibrium physics [29, 30].

A specific but important type of non-equilibrium states is realized when a stochastic system is coupled with external fields (via time-dependent, periodic transition rates) and a thermal bath. If the time-dependency is absent, the system settles into the stationary, equilibrium state $p_i^s = p_i^{eq.} \propto e^{-E_i/T}$. The detailed balance property (time-reversibility) implies that all probability currents between any pairs of states vanish. These features are absent when the external fields are time-periodic. The Floquet theory shows that the system settles into a time periodic state [31, 32], i.e. the probabilities of states become time-periodic functions with the same period as the external field. During one time-period the stochastic system jumps back and forth between its states, and this motivates us to look at features of time-integrated (over the period) probability fluxes from one state to another.

The physical implication of integrated probability current is seen in experiments on artificial molecular catenane systems [33], where rotation of smaller ring like molecule along the larger molecule is induced by external periodic perturbations. These molecules consist of two or more interlocked rings, the stationary larger ring and the smaller rings which can jump between binding sites around the larger one. These systems are called molecular rotors since

they are made up with stator (stationary larger ring) and rotators (smaller rings that can rotate around the larger ring). However, in order to use catenanes as molecular motors, i.e. to use it for performing work, unidirectional rotation of smaller rings is needed. The directed motion emerges solely due to an external fields, e.g. laser light and temperature changes result to cyclic conformational changes at binding sites which enables small rings to jump between discrete binding sites in unidirectional manner. It is obvious that without external fields thermally induced transitions will not create any unidirectional motion (clockwise or counterclockwise) on average.

The molecular motors is an area of intense research and there has been achievements in synthesizing artificial molecular motors. These artificial molecular machines are much smaller than the cell's size, so they can be used to move things at micro level, change the form, contract muscles, attack cancer cells *etc.* Artificial machines are not autonomous, they are controlled by external parameters such as temperature change, laser light, or chemical gradients. Many examples of these small nano-machines are presented in [34], which include molecular shuttles, switches, ATP, DNA-based machines, catenanes [33, 34] *etc.* Because of the small sizes, molecular systems display strong fluctuation, hence they are studied in the framework of stochastic processes. Understanding how to control these small motors by external fields in thermal environment raises new theoretical challenges. One of the control techniques is when external parameters are changed periodically which can pump a desired directed motion. Hence the term stochastic pump is introduced to describe periodically driven discrete-state Markov systems by examining probability currents in the periodic regime. This situation is met in the above mentioned catenane system where it was shown that in order to obtain directed motion one should also vary barriers between binding sites rather than only binding energies [33, 35]. This is the so-called no-pumping effect when probability currents between different states are zero.

Rahav, Horowitz and Jarzynski [36] first derived a no-pumping theorem (no-pumping theorem) which shows that the time-averaged probability currents nullify in the steady periodic (Floquet) state generated by time-periodic external fields acting on a Markov stochastic system when transitions between states have Arrhenius-type time-dependence [37]. This feature

of nullification makes an analogy between features of steady periodic and equilibrium states, because in the latter situation all probability currents vanish explicitly. Later works include those in [32,35,38–44]. However, the assumption on the Arrhenius rates is fairly specific, and it need not be met in applications. In this work a new mechanism is identified for the no-pumping theorem, which holds for symmetric time-periodic external fields and the so called destination rates (see 1.6). These rates are the ones that lead to the locally equilibrium form of the master equation, where dissipative effects are proportional to the difference between the actual probability and the equilibrium (Gibbsian) one. The mechanism also leads to an approximate no-pumping theorem for the Fokker-Planck (1.5) rates that relate to the discrete-space Fokker-Planck equation.

A system coupled with two or more thermal baths having different temperatures is another simple example of non-equilibrium situation, because there is a continuous energy flux through the system [45]. Since the system is much smaller than the baths (i.e. the surrounding environment), the system will settle into non-equilibrium steady state (NESS), where final bath temperatures are equal to the initial temperatures. In a NESS, the probabilities of states do not change in time, hence the properties of the system like energy, concentrations remain constant, but there are nonzero fluxes of energy, particles *etc* through the system. For example, proper concentrations of various ionic species inside cells, e.g. sodium, potassium, are maintained constant through continually moving ions across the cell's membrane [46]. Importantly, homeostatic processes which ensure a constant internal environment in cells lead to a NESS [47].

For a Markovian dynamics with time-independent transition probabilities, the Perron-Frobenius theorem ensures that system reaches a NESS [48]. Moreover, it is unique if the process is ergodic, i.e. there is a nonzero transition probabilities between each pair of states. In contrast to equilibrium, where the probability distribution over the states is given by the Boltzmann's distribution, there is no similar explicit formula for NESS. Such NESS probability distributions are computed numerically from the master equation. A stochastic description based on master equation provides a framework to non-equilibrium thermodynamics (stochastic thermodynamics) enabling to extend thermodynamic concepts (such as entropy

production, work, free energy *etc*) beyond equilibrium states. The master equation is used to study systems at NESS, e.g. chemical reaction network at NESS [49–51] *etc*; see [45,49,52,53] for general discussions of NESS and its biochemical applications.

One can distinguish non-equilibrium stationary states from equilibrium ones via probability currents: in NESS there are states characterized by non-zero probability flows between them [54]. Thus, a NESS is characterized not only by p_i^s, stationary probability found from master equation, but by pair of distributions: (p_i^s, J_{ij}), where $J_{ij} \neq 0$ is probability current from j to i [54]. In fact probability currents play a central role in non-equilibrium systems. These currents encode physical properties of the system, because energy, matter, information fluxes can be computed with J_{ij}; in this context see the above-mentioned discussions of the no-pumping theorem experimentally observed in catenanes.

The effectiveness of thermodynamics is that its concepts can be applied to non-equilibrium states. The response of the equilibrium system to a small external field is expressed by the fluctuation-dissipation (F-D) relations [55,56] which relates thermal fluctuations of the system from its average state with the energy dissipation, e.g. Einstein–Smoluchowski relation [57] for the Brownian motion. The F-D theorem is typically formulated by means of linear response function [56]. Various approaches have been proposed for extensions of F-D theorems for systems slightly perturbed from NESS where the notion of effective temperature (ET) is used [58,59].

The concept of ET is reviewed in [60] with applications to ideal and real gases, electromagnetic radiation, nuclear collisions, granular systems, glasses, sheared fluids, amorphous semiconductors and turbulent fluids. Ref. [60] also presents different definitions of ET including those based on F-D relations. Various biological systems in NESS are understood via the notion of ET, e.g. red-blood cells [61], self-propelled particles in an autonomous regime [62], where their collective behaviour can adapt to changing environment *etc*. This adaptivity property plays a central role in designing artificial systems [63]. Another intriguing example of a system in NESS described by ET is the "hot Brownian motion" [64–69], i.e. the motion of a laser-heated nano-particle. Here the NESS is established due to a different temperatures of the laser-driven particle versus the environment. Due to a time-scale separation hot Brown-

ian particles carry with them a hot halo, this provides several nanotechnological applications, among them one can mention “Photothermal Correlation Spectroscopy" (PhoCS) [70].

Time-scale separation means that a part of a system operates very fast compared to the rest. Hence the dynamics of the system is described in terms of slow variables due to eliminating the fast ones (adiabatic elimination) [71]. The adiabatic elimination works because the fast variables have reached their steady state for fixed slow variables. Physical examples of time-scale separations include the elimination of fast variables for atomic motion in laser light [72], the separation of slow and fast time scales associated with temperatures in the climate model [73], the derivation of the hydrodynamic equations from the Boltzmann approach *etc*.

The idea of time scale separation is also widely used in different areas of biology, including gene regulation, ligand-gated ion channels, enzyme kinetics, G-protein coupled receptors, *etc* [74]. It was first introduced in biochemical systems to describe enzyme kinetics [74, 75] in which free enzyme and the enzyme-substrate complex are fast variables, while the substrate and product were regarded as slow variables.

The time-scale separation has important thermodynamic consequences, e.g. the Helmholtz free energy can involve only slow variables [76], i.e. it does not depend on fast variables. Recall that free-energy difference $\triangle F$ define the quasi-static (isothermal) work: $\triangle F = \overline{W}$ (the line over W denotes an average over an ensemble of measurements of W), i.e. the free energy is a measure of the amount of usable energy that can do work at finite temperature. In contrast, if the system is driven from one equilibrium state to another by non-quasi-static external parameters, then $\overline{W} \geq \triangle F$. Jarzynski derived his famous equality [77, 78] directly relating non-equilibrium work and free energy difference. Another non-equilibrium relationship between work and free energy difference was put forward by Crooks [79]. The Jarzynski and Crooks relations are important for findng free-energy profiles of reaction coordinates [80–82].

An open problem is how to extend the notion of free energy to NESS. This extension is to be sought in terms of Markov processes through the master equation [45, 76, 83–85].

Here we aim at characterizing the NESS of a discrete-state, continuous-time Markov system with time-scale separation when put in contact with two reservoirs [85]. In our model the

driving protocol is such that external fields act only on the slow variable. This assumption is quite realistic because it is normally difficult to have a precise control on variables that move fast. In our model the transition probabilities between states of slow variable have the so-called Activation rate form or slow variables live in three-like structure. These three conditions

- time-scale separation between variables of the system
- partial controllability (external fields couple only with the slow variable)
- specific transition rates with an effective detailed balance

are needed for the existence of free energy in this non-equilibrium system. Here we have shown a number of interesting features, e.g. in contrast to its equilibrium counterpart, the non-equilibrium free energy can increase with temperature.

When the environment is divided into three parts: hot, cold and work reservoirs, then a system can operate as heat engine: at one stage the system is heated absorbing energy from hot reservoir, at another it is cooled delivering energy to the cold reservoir, and some amount of heat energy input is used to perform work. Reverse operation transforms the system into a refrigerator or heat pump. Heat engines ushered in the Industrial Revolution, examples of them in daily life include Internal combustion (e.g. steam engine, steam turbines at power stations) and External combustion (e.g. diesel and petrol engines in cars and airplanes) engines.

Nowadays nanotechnological techniques makes possible to design these engines at the nano-scale, which are strongly influenced by thermal fluctuations. Hence they are modeled within the framework of quantum and stochastic thermodynamics [58, 86–100]. Stochastic heat engines (SHE) are conventional analogs of heat engines where random behaviour of energy currents play important role in the work-extraction. Non-zero energy currents are associated with non-zero probability fluxes which define average energy lost or gained by thermal bath because for states j and i with energies E_j, E_i the probability flux $(\rho_{i\leftarrow j}p_j - \rho_{j\leftarrow i}p_i)$ indeed define energy lost(gained) through $(E_i - E_j)(\rho_{i\leftarrow j}p_j - \rho_{j\leftarrow i}p_i)$. With discrete number of states these energy flows is resulted from population inversion between the states of the system. Note that SHEs are also studied in NESS regime.

The creation of small engines is very appealing due to the need of energy harvesting [101, 102] i.e. extracting energies from ambient sources (e.g. solar power, wind energy, temperature gradients, tidal power *etc*). Capturing small amount of energies from ambient sources has also economical and ecological importance, e.g. it leads to burning less fossil fuel (hence produce less flue-gas such as CO_2) for generating electricity. Temperature differences are everywhere in both natural and human-made environment. Hence it can be used to do work, e.g. thermoelectric generators (TEG) convert heat (temperature gradient) into electrical current using Seebeck effect [103]; see the recent work [104] of present and future applications of TEGs. Note that the efficiency of TEGs like for any heat engine is limited by the Carnot's theorem.

Like TEGs SHEs can operate whenever there is temperature gradient. The origins of SHEs can be traced back to thought experiments of thermal (Brownian) ratchets, system consisting of a ratchet connected to a paddle wheel by an axle. The system is imagined to be small enough that the wheel can rotate bombarded by randomly moving molecules. The idea of a hypothetical ratchet mechanism was introduced around 1900s by French physicist Lipmann as perpetual motion machine of the second kind since since he believed that the ratchet will rotate in one direction only. In 1912, Smoluckowski first showed that there is no violation of the Second Law of thermodynamics [105]. The problem was later discussed and extended by Feynman [106, 107], in 1963, who showed that such a mechanism can work as a heat engine when different parts are at different temperatures. The Feynman's ratchet model fostered the development of Brownian motors [87, 108–112], nano-scale systems that do work in the presence of thermal noise and external fields. Fields can be in the form of time-periodic driving, ratchet-like (spatially asymmetric periodic) potentials, temperature, concentration or chemical gradients *etc*. Note that Brownian motors with asymmetric environment can result to a net pumping of particles into the some direction [108], e.g. motor proteins exhibit this phenomenon for intracellular transport. In this context, ratchet-based periodically driven molecular pumps [108] are very similar to the system discussed in section first chapter. Ref. [110] discussed a system driven by time-periodic changes of temperature, where the system is in the Floquet regime.

Experimentally, SHEs including the Brownian heat engines have been realized recently [87,113–117], their realization is mainly due to the development of optical trapping techniques [118,119].

However, a major limitations for heat engines is that their functioning demands external control to perform thermodynamic cycle, e.g. the Carnot cycle is implemented by the specific sequence of adiabatic and isothermal processes. The realization of Carnot cycle at the micro-scale with optically trapped Brownian particle is discussed in [113,114].

Smaller engines may be autonomous [96,97], but they are not adaptive to environmental changes, e.g. for fixed environment (thermal baths) there are internal parameters, under which the machine may act as a heat-pump or refrigerator. Hence they do demand fitting between internal and environmental parameters [89–95].

We study a model for an adaptive heat engine [120], where due to feedback from the functional part the engine's structure adapts to given thermal baths. Hence no on-line control and no external fitting are needed. Our approach is motivated by photosynthesis: the major heat engine of life that operates between the hot Sun temperature and the low-temperature Earth environment. It does have adaptive features that allow its functioning under decreased hot temperature (shadowing) or increased cold temperature (hot whether) [121,122]. Such engines can be useful for fuelling devices employing unknown and/or scarce resources; e.g. they can adapt to results of they own functioning that makes the bath temperatures closer.

Both in equilibrium and non-equilibrium stationary states one needs to compute averages of observables by summing or integrating over high-dimensional state space. In equilibrium situation the state space is determined by Boltzmann weights, hence averages of observables are computed by Boltzmann distribution. However, apart from a few models [123], it is not usually possible to compute analytically such averages or partition functions in statistical physics. At the same time the state space of physical systems is very large, e.g. in 3-dimensional Ising model the number of spin configurations with particles at 10^N lattice sites is 2^{10^N}, or for a system consisting of N classical particles one needs to compute $6N$ dimensional phase space integrals. In the absence of exact solutions numerical approaches are used. In any numerical approach the goal is to compute integrals with small number of function evaluations.

Hence the results are obtained at the the price of a statistical errors. Classical numerical integration techniques include Newton-Cotes type formulae, Gaussian quadratures, Romberg integration [124, 125] *etc.* However, classical numerical integration methods are well applied to one-dimensional cases [126].

The Monte Carlo(MC) method is one of the most important numerical techniques in statistical physics and not only [127, 128] which is mainly used for high-dimensional integration, random sampling and optimization [129]. The idea behind MC method is simple: to obtain numerical results by random sampling. Hence a central question is how to draw random samples according to a specified distribution without so much numerical efforts. In [130] Metropolis et. al. introduced an algorithm (Metropolis also known as Metropolis-Hastings algorithm) to generate random samples with desired distribution. The idea behind is to generate Markov chain over state space whose limiting distribution is of interest [1, 130–134]. In equilibrium situation this means to generate random states with Boltzmann weights. In order to ensure relaxation to equilibrium state transition probabilities can not be arbitrary. A sufficient (but not necessary) condition for transition rates that define stationary state is detailed balance constraint (1.8,3.2), which define stationary state: $dp_n/dt = 0$ (1.7,3.1). One of the choices of rates are Metroplis rates [130] (see the first section), the first choice used in statistical physics [135].

The central component of any MC method including Metropolis algorithm are random numbers, and it was their use that promoted the development of pseudo-random number generators(PRNG) [129], mathematical algorithms by which computers produce random numbers. By contrast, true random number generators (TRNG) produce random numbers from physical phenomena. However to use TRNGs in MC simulations is not practical since they are not fast in the sense of generating numbers. They are also non-deterministic(numbers can not be reproduced) and have no period. Most PRNGs produce random numbers which are uniformly distributed in the interval of [0,1]. Hence PRNGs should pass statistical tests of uniformity. While satisfying the criterions of "good" PRNG such as long period, portability, repeatability, efficiency, uniformity, however statistical properties and time characteristics of PRNGs are crucial to consider a generator as "good".

To address these challenges the renewed version of MIXMAX PRNG [136, 137] based on Anosov C-systems and Kolmogorov K-systems has been introduced in [2,138,139]. The MIXMAX is matrix-recursive PRNG and it has been shown that the properties of the MIXMAX generator is improved with increasing the size N of MIXMAX matrix [2].

More recently, the MIXMAX has been included in ROOT and Class Library for High Energy Physics (CLHEP) software packages [140] and claims to be a state ofart generator due to its long period, high performance and good statistical properties. In this paper the various statistical tests for MIXMAX are performed. The results compared with those obtained from other PRNGs, e.g. Mersenne Twister [3], Ranlux [4], Linear congruential generator (LCG) reveal better qualities for MIXMAX in generating random numbers. The Mersenne Twister is by far the most widely used PRNG in many software packages including packages inHigh Energy Physics (HEP), however the results show that MIXMAX is not inferior to Mersenne Twister [141].

Chapter 1

No-pumping theorem for non-Arrhenius rates [1]

1.1 Introduction

A wide range of systems appearing in physics, chemistry and biology can be modeled by Markov processes. Physically, Markov dynamics is the main tool for describing open systems (both quantum and classical) that interact with energy and/or particle reservoirs [143]. Hence it is at the core of non-equilibrium thermodynamics [142]. It is also the main tool for describing chemical reactions [9]. Among its biological applications one can mention conformational dynamics of biological molecules [144,145], ion channel gating processes [146], dynamics of predation, epidemic processes, genetics of inbreeding [147] *etc.* Such applications are frequently developed within random walk models, e.g. chemotaxis, biological motions [148] *etc.*

Generically, a Markov dynamics with time-independent transition rates relaxes to a stationary state. For a single-temperature reservoir (equilibrium thermal bath) this stationary state amounts to the Gibbs distribution at the bath's temperature [9, 143]. The equilibrium nature of the bath is reflected in the detailed balance condition that ensures nullification of

[1] The results considered in this chapter are published in Ref. [32].

all probability currents in equilibrium [9, 143].

The concept of the stationary state is generalized, if the stochastic system is subject to an external time-periodic field [31, 143, 149]. The system still forgets its initial conditions and appears in a non-equilibrium, time-dependent state, whose probabilities oscillate with the same period as the external field. This is the content of the Floquet theorem (outlined below), and this motivates us to look at features of time-integrated (over the period of the field) probability currents from one state to another. Now the no-pumping theorem [32, 35, 36, 38–41, 43, 44] states that probability currents nullify for an arbitrary time-periodic external field provided that the (time-dependent) transition rate $\rho_{i\leftarrow j}(t) > 0$ from state j to i holds the Arrhenius form [37]

$$\rho_{i\leftarrow j}(t) = e^{B_{ij}+\beta E_j(t)}, \tag{1.1}$$

where $B_{ij} = B_{ji}$ refers to the time-independent transition state, $\beta = 1/(k_{\mathrm{B}}T)$ is the inverse temperature, and $E_j(t)$ is the oscillating energy of the state j. Transitions from one state to another are induced by a thermal bath at temperature T, because if the bath is absent then due to energy conservation transitions between different states are also absent. Transition rates $\rho_{i\leftarrow j}(t)$ can be time-dependent solely due to an external field that acts on the system making its energies $E_j(t)$ time-dependent.

Thus the no-pumping theorem shows that the non-equilibrium, time-dependent state still holds an effectively equilibrium feature of nullifying (time-average) currents. (Whenever also B_{ij} in (1.1) are time-dependent, non-zero time-averaged currents are not excluded.) Hence the theorem fits naturally to the continuing effort of understanding the statistical mechanics of periodically driven systems using analogies with the equilibrium (i.e. time-independent) situation [149–152]. Recent works established several interesting relations between a driven system that hold the detailed balance condition and a similar system that is kept unnder constant (time-independent) non-equilibrium conditions [153]; in this context see also [154, 155].

Note that the same proof of the no-pumping theorem applies to rates more general than

(1.1), but this generalization (though useful for its own sake) is achieved at the cost of violating the detailed balance condition [39]. I.e. formally the results of [39] refer to non-equilibrium baths.

The virtue of the no-pumping theorem is that it applies to all oscillating external fields. Its major drawback is that the Arrhenius form (1.1) does not hold in many important applications, where simultaneously the detailed balance is required. For example, the Metropolis rates (the main tool of the Monte-Carlo dynamics), hold (1.1) with

$$B_{ij} = -\beta \max[E_j(t), E_i(t)] \tag{1.2}$$

hence

$$\rho_{i\leftarrow j}(t) = \min[1, e^{\beta[E_j(t)-E_i(t)]}]) \tag{1.3}$$

We see that B_{ij} cannot stay time-independent, if $E_i(t)$ and $E_j(t)$ are time-dependent. Further important examples of non-Arrhenius rates include Kramers rates that emerge out of diffusion in energy landscape [9] and corresponds in (1.1) to

$$\rho_{i\leftarrow j}(t) = e^{-\beta\delta_{ij}+\beta(\,E_j(t)-\max[E_i(t),E_j(t)]\,)}, \tag{1.4}$$

where $\delta_{ij} = \delta_{ji}$ is energy barrier or activation energy that separates E_i and E_j. Another important example is the Fokker-Planck rates

$$\rho_{i\leftarrow j}(t) = e^{\beta[\,E_j(t)-E_i(t)\,]/2}, \tag{1.5}$$

which allow to match the discrete-space master equation for the Markov dynamics with the continuous-space Fokker-Planck equation [156]. For all these cases, the standard formulation of the no-pumping theorem would just allow non-zero time-averaged currents for a suitably chosen external field, i.e. the theorem is not very informative.

Here the aim is to extend the no-pumping theorem to rates different from (1.1) (the detailed balance is always assumed to hold).

First, it will be shown that the no-pumping theorem|time-integrated probability currents nullify|holds for the destination rates

$$\rho_{i\leftarrow j}(t) = e^{-\beta E_i(t)}, \tag{1.6}$$

under an additional sufficient condition that the external fields are (effectively) time-symmetric. The mechanism is more general since it nullifies the currents for certain non-symmetric external fields as well.

The destination rates (1.6) lead to the locally-equilibrium form of the master equation, which is driven by the difference between the actual probability and the equilibrium one; see the discussion after Eq. (1.23). There is a long and successful tradition of applying locally-equilibrium master equations in non-equilibrium physics. It was initiated via the model proposed in 1954 by Bhatnager, Gross and Krook [157–161], and since that time proved to be very useful [162]. In particular, the rates (1.6) were employed in [163] for describing the dynamics of a paradigmatic disordered statistical systems (the Random Energy Model), and found to be in agreement with experiments. Below I show that|in contrast to the Arrhenius rates (1.1)|the destination rates provide a reasonable approximation for other rates (e.g. the Fokker-Planck rate). Hence their experimental success is not accidental.

Second, it will be demonstrated numerically that the same mechanism that leads to the exact no-pumping theorem for the destination rates ensures an approximate validity of this theorem for the Fokker-Planck rate.

1.2 Master-equation and the Floquet theorem

Let a system can be in discrete states $[i = 1, ..., n,]$. A master equation describes the time evolution of the probability of a system to occupy the state i

$$\dot{p}_i \equiv dp_i/dt = \sum\nolimits_j [\rho_{i\leftarrow j}(t)p_j - \rho_{j\leftarrow i}(t)p_i], \tag{1.7}$$

where $p_i(t)$ is the probability of a system to be in state i at time t, and $\rho_{i\leftarrow j}(t) > 0$ is the transition rate from j to i. It is assumed that for any fixed time t, there is the global detailed balance at inverse (time-independent) temperature β:

$$\rho_{i\leftarrow j}(t)\, e^{-\beta E_j(t)} = \rho_{j\leftarrow i}(t)\, e^{-\beta E_i(t)}. \tag{1.8}$$

Note that $\rho_{i\leftarrow j}(t)$ are rates, not probabilities, hence they have units of $[time]^{-1}$, and they are not restricted to the interval $[0, 1]$. The master equation conserves the total probability $\frac{d}{dt}\sum_i p_i = 0$ so that the normalization is preserved with a normalized initialization. Moreover, master equation with positive rates ensures positive probabilities $p_i \geq 0$. Eq. (1.7) can be written in the matrix form

$$\dot{p}_i = \sum\nolimits_{j=1}^{n} w_{ij} p_j, \tag{1.9}$$

where w_{ij} matrix elements are defined as follows

$$\begin{aligned} w_{ij} &= \rho_{i\leftarrow j} \qquad \mathrm{i} \neq \mathrm{j}, \\ w_{ii} &= -\sum\nolimits_{j\neq i} \rho_{j\leftarrow i} \end{aligned} \tag{1.10}$$

Due to external field(s) acting on the system, the energies $E_i(t)$ are time-periodic functions with period τ:

$$E_i(t) = E_i(t+\tau). \tag{1.11}$$

The instantaneous probability flux from state j to state i is

$$J_{ij}(t) = \rho_{i\leftarrow j}(t) p_j(t) - \rho_{j\leftarrow i}(t) p_i(t), \quad J_{ij} = -J_{ji} \tag{1.12}$$

Before specifying the external field, let me remind the Floquet theorem, which is necessary for defining the no-pumping theorem. Using the normalization of probabilities $\sum_{i=1}^{n} p_i = 1$, we obtain from (1.9)

$$\dot{p}_i = \sum\nolimits_{j=1}^{n-1} w_{ij} p_j + w_{in}(1 - p_1 - \ldots - p_{n-1}) = \sum\nolimits_{j=1}^{n-1} (w_{ij} - w_{in}) p_j + w_{in}, \tag{1.13}$$

Thus we get from (1.7, 1.9, 1.13)

$$\dot{P} = W(t)P(t) + b(t) \tag{1.14}$$

where $P(t) = [p_1(t), ..., p_{n-1}(t)]$ and $b(t)$ are $(n-1) \times 1$ vectors and $W(t)$ is $(n-1) \times (n-1)$ matrix:

$$b_i = W_{in}, \quad W_{ij} = w_{ij} - w_{in}, \ i, j = 1, \ldots, n-1, \tag{1.15}$$

$$w_{ij} = \rho_{i \leftarrow j} - \delta_{ij} \sum\nolimits_{j \neq i} \rho_{j \leftarrow i}, \quad i, j = 1, \ \ldots, \ n. \tag{1.16}$$

The solution of (1.14) with initial condition $P(t_0)$ is

$$\begin{aligned} P(t) &= A(t, t_0)P(t_0) \\ &+ \int_{t_0}^{t} \mathrm{d}s \, A(t, s)b(s), \quad A(t, s) \equiv \overleftarrow{e}^{\int_s^t du W(u)}, \end{aligned} \tag{1.17}$$

where $\overleftarrow{e}$ is time-ordered or chronological exponent. For $t \gg t_0$, the state $P(t_0)$ is forgotten, which is equivalent to $A(t, t_0)P(t_0) \to 0$. Taking $t_0 = -\infty$ in (1.17) we get from (1.17)

$$P(t) = \int_{-\infty}^{t} \mathrm{d}s \, A(t, s)b(s). \tag{1.18}$$

Recalling that $W(t)$ and $b(t)$ are time-periodic with the same period, by making the substitution $s - \tau = x$ in (1.18) we see that $P(t)$ is also time-periodic with the same period.

$$P(t + \tau) = P(t) \tag{1.19}$$

This is the content of the Floquet theorem: for sufficiently long times, the stochastic system subject to time-periodic driving appears in a steady periodic state. This motivates us to characterize this state via time-averaged probability currents [cf. (1.12)]

$$\Phi_{ij} = \frac{1}{\tau} \int_{a}^{a+\tau} J_{ij}(t) dt, \tag{1.20}$$

where once the system is in its steady periodic (Floquet) state due to $a \gg t_0$, Φ_{ij} does not anymore depend on a. Below I shall employ $a = 0$ in the averaging. This implies that initial conditions are posed at much earlier time: $t_0 \to -\infty$. From antisymmetry of J_{ij}s (1.12) it follows that the integrated currents are also antisymmetric:

$$\Phi_{ij} = -\Phi_{ji} \tag{1.21}$$

Note that we do not consider cases, where the transition matrix describes an reducible chain. There the system does not generally forget its initial state.

1.3 No-pumping theorem for a single field and destination rates

Using the normalization condition for probabilities

$$\sum_j p_j(t) = 1 \tag{1.22}$$

we obtain from (1.7) for the destination rates (1.6)

$$\dot{p}_i = e^{-\beta E_i(t)} - p_i(t) Z(t), \quad Z(t) \equiv \sum\nolimits_j e^{-\beta E_j(t)}. \tag{1.23}$$

This is the main advantage of rates (1.6): the equations for the probabilities decouple from each other, making it convenient for studying systems with irregular distribution of energies [163]. Note that (1.23) can be written as

$$\dot{p}_i = -Z(t)[\, p_i(t) - \frac{e^{-\beta E_i(t)}}{Z(t)}\,] \tag{1.24}$$

showing that the change $\dot{p}_i$ of the probability is proportional to the difference between this probability and its equilibrium value $\frac{e^{-\beta E_i(t)}}{Z(t)}$. This makes connection between the studied destination rates and the Bhatnager, Gross and Krook kinetic equation [157–162].

Now (1.6, 1.12) imply

$$J_{ij} = \dot{p}_i p_j - \dot{p}_j p_i = p_j e^{-\beta E_i} - p_i e^{-\beta E_j}. \tag{1.25}$$

The no-pumping statement I propose is that for field (1.11) [plus additional symmetry conditions to be specified below], and for rates (1.6), it holds

$$\langle \dot{p}_i p_j \rangle \equiv \int_0^\tau \mathrm{d}t\, \dot{p}_i(t) p_j(t) = 0, \tag{1.26}$$

thereby nullifying also the time-averaged current $\Phi_{ij} = 0$; see (1.25, 1.20). Note (1.26, 1.23) can be written as

$$\langle e^{-\beta E_i} p_j \rangle = \langle p_i p_j Z \rangle. \tag{1.27}$$

Hence the validity of (1.26) is obvious in the limiting case of very slow time-dependence, where the probabilities freeze to their Gibbsian (quasi-equilibrium) values: $p_i(t) = e^{-\beta E_i(t)}/Z(t)$.

To prove (1.26), we start from (1.23) and introduce there a new time-variable s

$$\frac{\mathrm{d}s}{\mathrm{d}t} = Z(t), \quad s = \int_0^t \mathrm{d}u\, Z(u). \tag{1.28}$$

Due to $Z(t) > 0$, the s-time relates to the t-time by a one-to-one mapping. Since $Z(t+\tau) = Z(t)$ [see (1.23, 1.11)], we get from (1.28):

$$s(t+\tau) = s(t) + \sigma, \quad \sigma = \int_0^\tau \mathrm{d}u\, Z(u). \tag{1.29}$$

Thus if $p_i(t)$ (in the Floquet regime) is τ-periodic, $p_i(t) = p_i(t+\tau)$, then $p_i(s)$ is σ-periodic:

$$p_i(s) = p_i(s+\sigma). \tag{1.30}$$

Note that the integral in (1.26) stays invariant under changing the time:

$$\int_0^\tau \mathrm{d}t\, \dot{p}_i(t) p_j(t) = \int_0^\sigma \mathrm{d}s\, \frac{\mathrm{d}p_i(s)}{\mathrm{d}s}\, p_j(s). \tag{1.31}$$

We get from (1.23)

$$\frac{\mathrm{d}p_i(s)}{\mathrm{d}s} = -p_i(s) + e^{-\beta E_i(s)}/Z(s). \tag{1.32}$$

We now introduce the Fourier-expansion for σ-periodic functions $g(s+\sigma) = g(s)$

$$g(s) = \sum_{n=-\infty}^{\infty} \hat{g}_n e^{\frac{2\pi \mathrm{i} s n}{\sigma}}, \tag{1.33}$$

$$\hat{g}_n = \int_0^\sigma \frac{\mathrm{d}s}{\sigma}\, g(s)\, e^{-\frac{2\pi \mathrm{i} s n}{\sigma}} = \int_{-\sigma}^{\sigma} \frac{\mathrm{d}s}{2\sigma}\, g(s)\, e^{-\frac{2\pi \mathrm{i} s n}{\sigma}}. \tag{1.34}$$

and apply it to $p_i(s) \to \hat{p}_{i,n}$ and $e^{-\beta E_i(s)}/Z(s) \to \hat{\psi}_{i,n}$. Note that $\hat{g}_n^* = \hat{g}_{-n}$, since $g(s)$ is real. Substituting Eq. (1.33) into Eq. (1.32) we obtain

$$\sum_{n=-\infty}^{\infty} \frac{2\pi \mathrm{i} n}{\sigma} \hat{p}_{i,n} e^{\frac{2\pi \mathrm{i} s n}{\sigma}} = -\sum_{n=-\infty}^{\infty} \hat{p}_{i,n} e^{\frac{2\pi \mathrm{i} s n}{\sigma}} + \sum_{n=-\infty}^{\infty} \hat{\psi}_{i,n} e^{\frac{2\pi \mathrm{i} s n}{\sigma}}, \tag{1.35}$$

which implies the following relationship between $\hat{p}_{i,n}$ and $\hat{\psi}_{i,n}$

$$\hat{p}_{i,n} = \hat{\psi}_{i,n} \left[1 + \frac{2\pi \mathrm{i} n}{\sigma}\right]^{-1}. \tag{1.36}$$

Using (1.36) we obtain for the integral in (1.31)

$$\int_0^\sigma \mathrm{d}s\, \frac{\mathrm{d}p_i(s)}{\mathrm{d}s}\, p_j(s) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \frac{2\pi \mathrm{i}\, n}{\sigma} \hat{p}_{i,n} \hat{p}_{j,m} \int_0^\sigma \mathrm{d}s\, e^{\frac{2\pi \mathrm{i} s (n+m)}{\sigma}}, \tag{1.37}$$

where we employed the Fourier expansion of (1.32). Now keeping the terms $m = -n$

$$\begin{aligned} \int_0^\sigma \mathrm{d}s\, \frac{\mathrm{d}p_i(s)}{\mathrm{d}s}\, p_j(s) &= \sum_{n=-\infty}^{\infty} \frac{2\pi \mathrm{i}\, n}{\sigma} \hat{p}_{i,n} \hat{p}_{j,-n} \int_0^\sigma \mathrm{d}s \\ = \sum_{n=-\infty}^{\infty} \frac{2\pi \mathrm{i}\, n\, \hat{\psi}_{i,n} \hat{\psi}_{j,-n}}{1 + (\frac{2\pi n}{\sigma})^2} &= -\sum_{n=1}^{\infty} \frac{4\pi\, n\, \mathrm{Im}[\hat{\psi}_{i,n} \hat{\psi}_{j,-n}]}{1 + (\frac{2\pi n}{\sigma})^2}, \end{aligned} \tag{1.38}$$

where we used relation (1.36). If now $\hat{\psi}_{i,n}\hat{\psi}_{j,-n} = \hat{\psi}_{i,n}\hat{\psi}_{j,n}^*$ is real, the sum in (1.38) is zero. Thus the integrals in (1.38, 1.31, 1.26) nullify.

Let now $E_i(t)$ in (1.11) are even:

$$E_i(t) = E_i(-t). \tag{1.39}$$

Then $s(t)$ is an odd function of t [see (1.28,1.29)], and hence $E_i(s) = E_i(-s)$. Then $\hat{\psi}_{i,\,n}$ and $\hat{\psi}^*_{j,\,n}$ are real [see (1.34)], and the integral in (1.38, 1.31, 1.26) nullifies thereby proving the no-pumping theorem. A more general situation, when the same reasoning applies, and $\hat{\psi}_{i,\,n}\hat{\psi}_{j,\,-n}$ is real, takes place when $E_i(t)$ in (1.11) can be made even after a suitable time-shift γ which does not depend on i:

$$E_i(t-\gamma) = E_i(-t-\gamma). \tag{1.40}$$

This is because in the Floquet regime the origin of time can be chosen arbitrary. In particular, (1.40) includes (1.46), where φ_j does not depend on j.

To prove that (1.40) makes $\hat{\psi}_{i,\,n}\hat{\psi}_{j,\,-n} = \hat{\psi}_{i,\,n}\hat{\psi}^*_{j,\,n}$ real, note that the new Fourier coefficient of shifted function is equal to the Fourier coefficient of unshifted function multiplied by the phase factor

$$\hat{\psi}^{new}_{i,\,n} = \frac{1}{\sigma}\int_0^\sigma \mathrm{d}s\,\frac{e^{-\beta E_i(s-\gamma)}}{Z(s-\gamma)}\,e^{-\frac{2\pi \mathrm{i} s n}{\sigma}} = \hat{\psi}_{i,\,n}e^{-\mathrm{i}\frac{2\pi n\gamma}{\sigma}}$$

Since $\hat{\psi}^{new}_{i,\,n}$ is real due to (1.40) it follows

$$\hat{\psi}_{i,\,n}\hat{\psi}_{j,\,-n} = \hat{\psi}^{new}_{i,\,n}\hat{\psi}^{new}_{j,\,n} = Real \tag{1.41}$$

I stress that (1.40) gives only a sufficient condition for the validity of (1.26). The following example illustrates this fact. Let me take $i = 1, 2, 3 = n$ (three-level system), $\sigma = 1$ and define energies $E_i(s)$ so that the following relations hold

$$e^{-\beta E_i(s)}/Z(s) = c_i + d_i s + f_i s^2 \quad \text{for} \quad 0 \le s \le 1, \tag{1.42}$$

while for $s > 1$ and $s < 0$, $e^{-\beta E_i(s)}/Z(s)$ is continued from (1.42) periodically with the period $\sigma = 1$. These functions are not continuous, but they can be considered as limits of continuous

functions. This suffices for the sake of the present example.

In (1.42), c_i, d_i and f_i are constants, which should ensure the normalization and positivity of the probabilities $e^{-\beta E_i(s)}/Z(s)$. In particular, I choose $\sum_{i=1}^{3} c_i = 1$ and $\sum_{i=1}^{3} d_i = \sum_{i=1}^{3} f_i = 0$ for normalization. Now generically (1.42) do not define any symmetric functions of s. However, we get

$$\hat{\psi}_{k,\,n} = \frac{f_k + \mathrm{i}\,(d_k + f_k)n\pi}{2n^2\pi^2}, \tag{1.43}$$

$$\mathrm{Im}[\hat{\psi}_{i,\,n}\hat{\psi}_{j,\,-n}] = \frac{f_j d_i - f_i d_j}{4\pi^3 n^3}. \tag{1.44}$$

The nullification of all currents amounts to $\mathrm{Im}[\hat{\psi}_{i,\,n}\hat{\psi}_{j,\,-n}] = 0$ for all i and j. Generally, this requires three conditions $f_j d_i = f_i d_j$ to be imposed on f_i and d_i. But due to $\sum_{i=1}^{3} d_i = \sum_{i=1}^{3} f_i = 0$, it suffices to take a single condition $f_1 d_2 = f_2 d_1$. This ensures $f_j d_i = f_i d_j$ and thus nullifies all currents.

1.4 Approximate no-pumping

The above no-pumping theorem concerns the destination rates (1.6). It is not valid exactly for other interesting rates, e.g. Kramers (1.4) or Fokker-Planck (1.5); see Figs. 1.1–1.3.

Now Figs. 1.1–1.3 show numerical results, where the time-averaged current for a three-level system is compared for three different rates: destination (1.6), Kramers (1.4) and Fokker-Planck (1.5).

Note that for a tree-level system(with connected states: $\rho_{j\leftarrow i} \geq 0, \forall i, j$) from Eqs. (1.21,1.7,1.19) in the periodic steady state we have [40]

$$\Phi = \Phi_{12} = \Phi_{23} = \Phi_{31}. \tag{1.45}$$

Numerics was carried out for the following concrete form of $E_i(t)$:

$$E_i(t) = \varepsilon_i + a_i \cos\left(\frac{2\pi t}{\tau} + \varphi_i\right), \quad i = 1, 2, 3, \tag{1.46}$$

where ε_i, a_i and φ_i are constants.

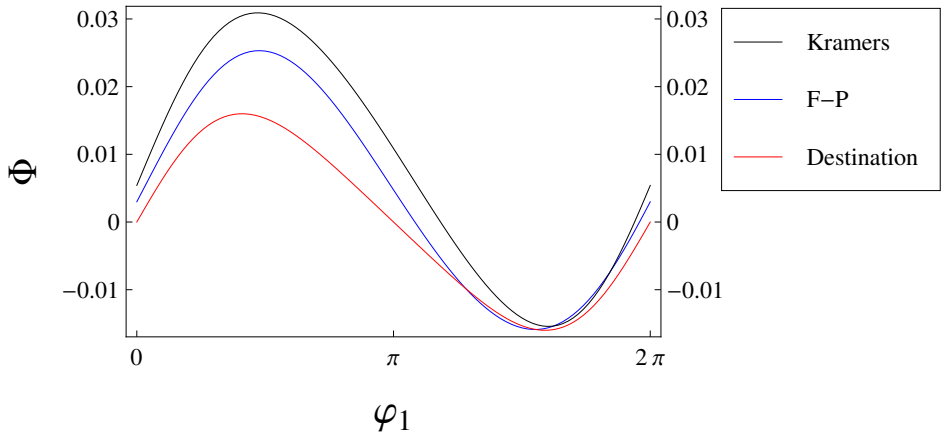

Figure 1.1: Time-averaged current $\Phi = \Phi_{12} = \Phi_{23} = \Phi_{31}$ given by (1.20) for a three-level system ($n = 3$) and $\beta = 1$ versus the parameter φ_1 for Kramers, Fokker-Planck (F–P) and destination rates; see (1.46). $E_i(t)$ are given by (1.46), where $\tau = 3$. Other parameters in (1.46): $\varphi_2 = \varphi_3 = 0$, and $\varepsilon_1 = \frac{1}{3}, a_1 = 1$, $\varepsilon_2 = \frac{2}{3}, a_2 = 2$, $\varepsilon_3 = 1, a_3 = 3$. For Kramers rates $\delta_{ij} = 1$ in (1.4).
Hence for $\varphi_1 = 0$ or $\varphi_1 = 2\pi$, the external field satisfies (1.40) and holds the no-pumping theorem $\Phi = 0$ for the destination rates, as seen on the figure. If (1.40) holds, $\Phi \approx 0$ for the Kramers rates (1.1, 1.4) and the Fokker-Planck rates (1.1, 1.5).

For (1.46) conditions (1.40) are satisfied e.g. for $\varphi_i = \varphi$ for all $i = 1, 2, 3$. This situation includes, e.g. the dipole coupling with an external, periodic electric field [143].

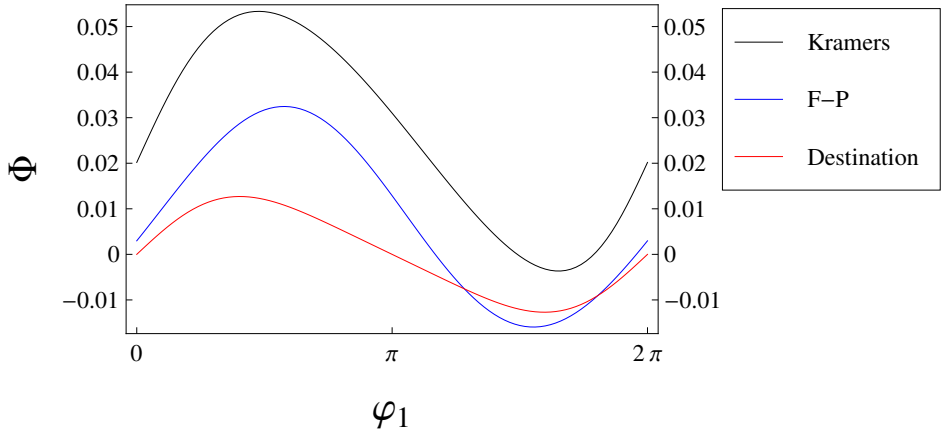

Figure 1.2: The same as in Fig. 1.1, but with $\tau = 1$, i.e. the external fields change faster than in Fig. 1.1, where $\tau = 3$. For this range of parameters the Fokker-Planck rates hold an approximate no-pumping theorem, while the Kramers rates do not.

Figures 1.1–1.3 refer to different values of the time-period τ in (1.46). Figs. 1.1–1.3 demonstrate that under condition (1.40), the value of the time-averaged probability current

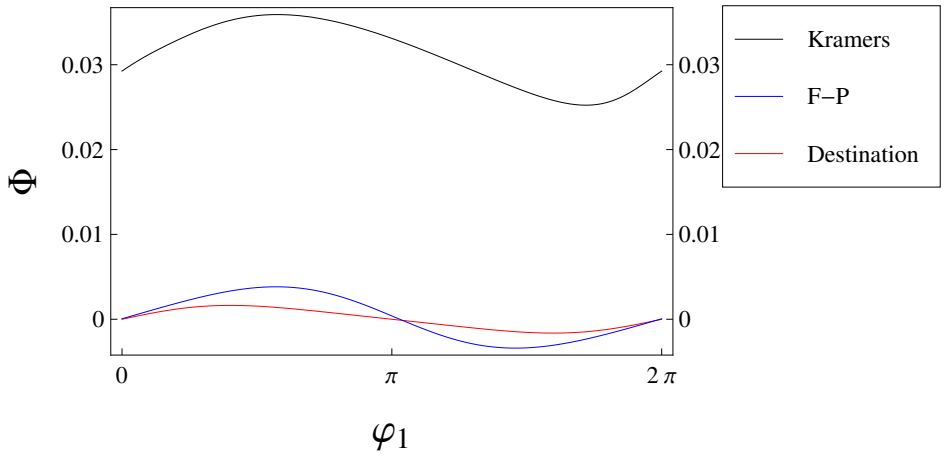

Figure 1.3: The same as in Fig. 1.1, but with $\tau = 0.1$; cf. Fig. 1.2. The external fields change faster than in Fig. 1.1 and in Fig. 1.3. The no-pumping theorem approximately holds for the Fokker-Planck (FP) rates.

nullifies exactly for the destination rates and it is approximately zero (with a good precision) for the Fokker-Planck rates (denoted as F–P in Figs. 1.1–1.3). For the Kramers rates the situation is different: it also predicts an approximately zero time-averaged probability current, but only for a sufficiently large τ; see Fig. 1.1.

Fig. 1.4 gives an example of a situation, where (for all studied rates) the time-averaged currents are sizable, since conditions (1.40) do not hold.

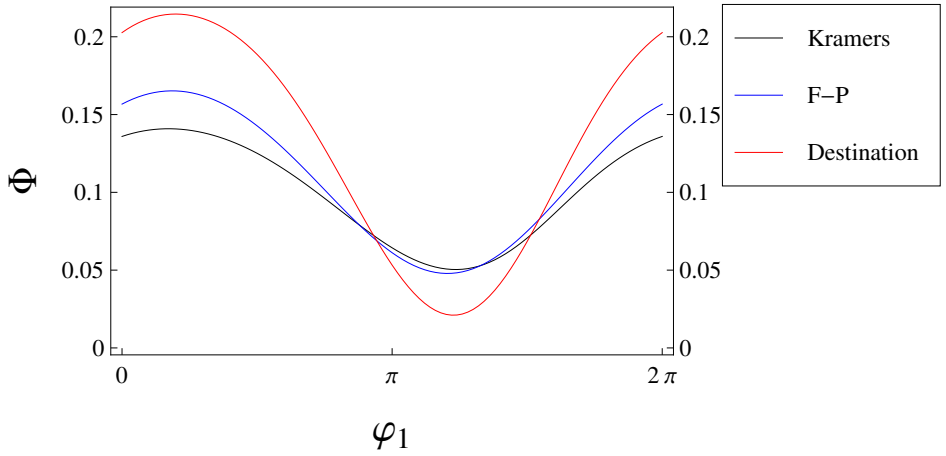

Figure 1.4: The same as in Fig. 1.1, but $\varphi_2 = \pi, \varphi_3 = \frac{3\pi}{2}$. In this example conditions (1.40) do not hold and the probability currents are sizable for all studied rates.

Note that all above numerical examples did not refer to high temperatures. Clearly, probability currents generally nullify for large temperatures, but one can identify a regime,

where the *instantaneous* time-dependent currents are still sizable, though their time-averages are practically zero. This is shown in Fig. 1.5, where the ratio between instantaneous and averaged currents amounts to $\sim 10^{-3}$. This high-temperature version of the no-pumping theorem holds for all studied rates and it does not need conditions (1.40).

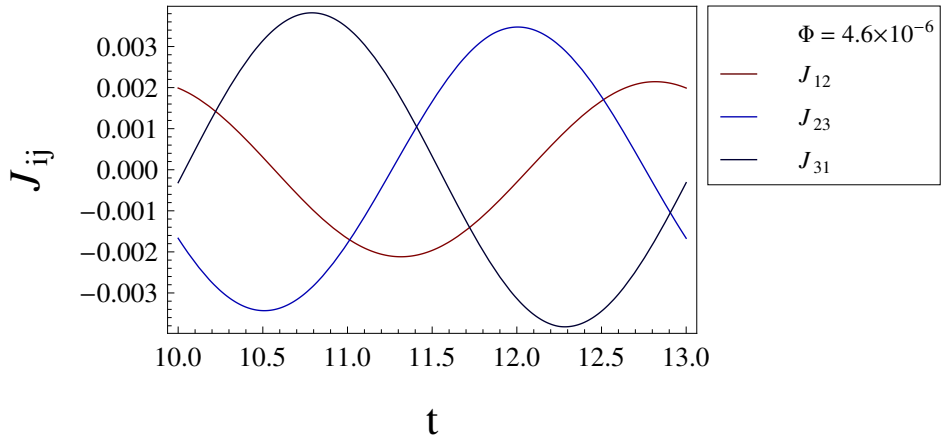

Figure 1.5: Instantaneous probability currents $J_{ij}(t)$ given by (1.12) for $\beta = 0.01$ and Kramers rates (1.1, 1.4). In (1.46) I took: $E_i(t) = -\frac{i}{2} + \frac{i}{2}\cos\left(\frac{2\pi t}{3} + \frac{\pi i}{2}\right)$, and for barriers: $\delta_{ij} = 1$. It is seen that J_{ij} are much larger than their time-average $\Phi = \Phi_{12} = \Phi_{23} = \Phi_{31}$.

1.5 Work

To keep the system in the non-equilibrium state, the external field dissipates work into the thermal bath. Now the work relates to energy (and not probability) currents through the system. Hence it is important to study it in the context of the no-pumping theorem.

The rate of work can be calculated via the standard formula

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \sum_i p_i(t)\dot{E}_i. \tag{1.47}$$

The positivity of the time-averaged work W is deduced from the positivity of the entropy production (see e.g. [58] for this concept, its physical meaning is clarified below)

$$S_e(t) = \frac{1}{2}\sum_{ik}(w_{ki}p_i - w_{ik}p_k)\ln\frac{w_{ki}p_i}{w_{ik}p_k} \geq 0, \tag{1.48}$$

where w_{ik} is defined in (1.16). Now writing as $\ln\frac{w_{ki}p_i}{w_{ik}p_k} = \ln\frac{w_{ki}}{w_{ik}} + \ln\frac{p_i}{p_k}$, we note that the second

term amounts in (1.48) to $-\frac{\mathrm{d}}{\mathrm{d}t}\sum_i p_i \ln p_i$

$$\begin{aligned}\frac{1}{2}\sum_{ik}(w_{ki}p_i - w_{ik}p_k)\ln\frac{p_i}{p_k} &= -\frac{1}{2}\sum_i \ln p_i \sum_k (w_{ik}p_k - w_{ki}p_i)\\ -\frac{1}{2}\sum_k \ln p_k \sum_i (w_{ki}p_i - w_{ik}p_k) &= -\frac{1}{2}\sum_i \dot{p}_i \ln p_i - -\frac{1}{2}\sum_k \dot{p}_k \ln p_k\\ &= -\sum_i \dot{p}_i \ln p_i = -\frac{\mathrm{d}}{\mathrm{d}t}\sum_i p_i \ln p_i, \end{aligned} \tag{1.49}$$

where we used Eq. (1.7). Due to Floquet theorem this term disappears after the time-averaging Eq. (1.48).

$$\int_0^\tau \mathrm{d}t\, S_e(t) = \frac{1}{2}\int_0^\tau \mathrm{d}t \sum_{ik}(w_{ki}p_i - w_{ik}p_k)\ln\frac{w_{ki}}{w_{ik}}. \tag{1.50}$$

From detailed balance condition (1.8) we have

$$\ln\frac{w_{ki}}{w_{ik}} = \beta E_i - \beta E_k. \tag{1.51}$$

Substituting (1.51) into (1.50) and following the same procedure as in (1.49) we get

$$\int_0^\tau \mathrm{d}t\, S_e(t) = -\beta\int_0^\tau \mathrm{d}t \sum_i \dot{p}_i E_i. \tag{1.52}$$

Thus we obtain from the first term the positivity of the time-averaged work:

$$\begin{aligned} W = \int_0^\tau \mathrm{d}t \sum_i p_i(t)\dot{E}_i &= -\int_0^\tau \mathrm{d}t \sum_i \dot{p}_i(t)E_i\\ &= T\int_0^\tau \mathrm{d}t\, S_e(t) \geq 0. \end{aligned} \tag{1.53}$$

Applying the Clausius inequality to the bath|recall that W turns to the heat Q received by the equilibrium thermal bath at temperature T, and then Q/T is smaller or equal to the bath entropy increase|it is seen that $\int_0^\tau \mathrm{d}t\, S_e(t)$ gives a lower bound for the bath entropy increase per cycle.

Note from Fig. 1.6 that the average work decays to zero both for high and low tempera-

tures. There is no no-pumping (i.e. no-work-dissipation) theorem for it.

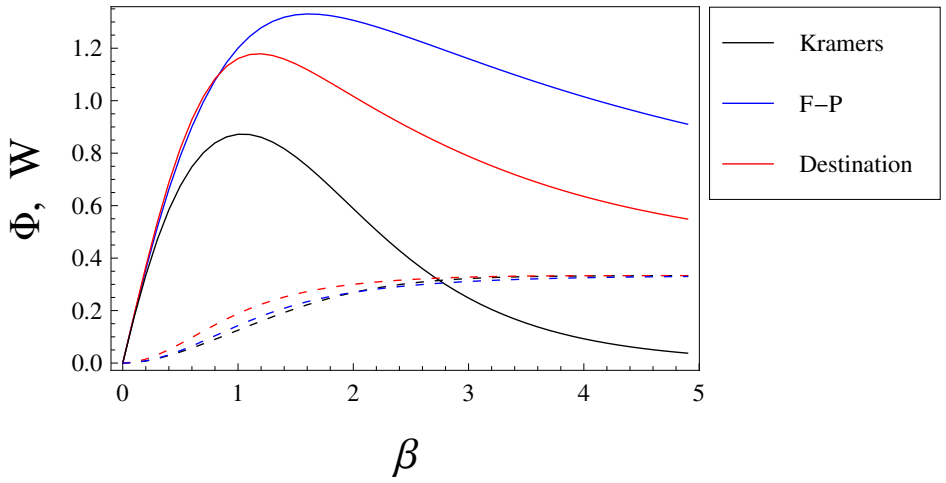

Figure 1.6: Time-averaged current $\Phi = \Phi_{12} = \Phi_{23} = \Phi_{31}$ (dashed curves) given by (1.20) and W given by (1.53) (full curves) versus the inverse temperature β and for various rates. In (1.46) I took: $E_i(t) = \frac{i}{3} + i\cos\left(\frac{2\pi t}{3} + \frac{i\pi}{2}\right)$.

Chapter 2

Free energy for non-equilibrium quasi-stationary states [1]

2.1 Introduction

One reason for effectiveness of thermodynamics is that its concepts and ideas apply beyond the domain of equilibrium states, e.g. in real life thermodynamics is applied to systems having different temperatures, even though such states do not belong to equilibrium. There are specific mechanisms for this applicability, e.g. thermodynamics applies to systems that are perturbatively close to equilibrium [142, 165–167]. Another general mechanisms is the time-scale separation. In fact, this mechanism is inherent in the structure of thermodynamics, which is built up via the notion of a quasi-static process [166]. It is also important in statistical physics of open systems, where even low-dimensional systems can play the role of a thermal bath, provided that they are fast [168, 169]. A related point is seen in stochastic thermodynamics, where the notion of white (i.e. fast) noise is relevant [170].

There is already a body of work concerning thermodynamic aspects of systems with time-scale separation [171–189]. In particular, much attention was devoted to the effective temperature of glassy systems, where the time-scale separation emerges from the many-body

[1] The results considered in this chapter are published in Ref. [85].

physics [171, 173–175]. Recent works studied time-scale separation in stochastic thermodynamics focusing on dissipative features such as entropy production [176–181].

Time-scale separation is a much wider notion, and it is frequent also in biology and society, where it allows to reduce the complexity of emergent structures [190,191]. Indeed, components of such systems enjoy a certain autonomy (though in different ways): fast variables "live" under fixed values of the slow ones, while the slow variables "see" stationary distributions of the fast ones.

We aim to look at a class of stationary, non-equilibrium states for a Markov stochastic system that is out of equilibrium due to interaction with two different thermal baths; see [45] for an introduction to such systems. We impose three conditions.

– Time-scale separation: the system under consideration consists of two variables, fast and slow.

– Partial controllability: external fields act on the slow variable only. This assumption can be validated from operational reasons: it is normally difficult to have a precise control on variables that move fast.

Both conditions need not hold for the equilibrium situation (equal temperatures of the thermal baths), where thermodynamics applies under any type of controllability and the ratio of characteristic times.

– Transition rates of the slow variable hold certain constraints|they are given by activation rates, or the slow variable lives in a tree-like structure|that amount to an effective detailed balance that holds after averaging (tracing out) over the fast variable.

Under these conditions the slow (isothermal) work done on this system admits a potential, i.e. there exists the free energy. This definition is unambiguous, because the work is defined for an arbitrary non-equilibrium state [167]; see the discussion after (2.15) for details of this point. Hence the free energy need not have further equilibrium features. Indeed, we show that its behavior with respect to temperature changes can be different from the features of the equilibrium free energy.

2.2 Two-temperature Markov dynamics

Consider the following master equation for two discrete variables $i = 1, ..., n$ and $\alpha = 1, ..., N$ (see e.g. [142, 170]):

$$\begin{aligned} \dot{p}_{i\alpha} = &\sum\nolimits_j [\rho_{ij|\alpha}\, p_{j\alpha} - \rho_{ji|\alpha}\, p_{i\alpha}] \\ &+\epsilon \sum\nolimits_\gamma [\omega_{\alpha\gamma|i}\, p_{i\gamma} - \omega_{\gamma\alpha|i}\, p_{i\alpha}], \end{aligned} \tag{2.1}$$

where $p_{i\alpha}$ is the joint probability of i and α, $\rho_{ij|\alpha}$ and $\omega_{\alpha\gamma|i}$ are the transition probabilities for $i \leftarrow j$ (for a fixed α) and $\alpha \leftarrow \gamma$ (for a fixed i), respectively. In (2.1), ϵ is a small parameter that makes $\{\alpha\}$ slower than $\{i\}$. We assume that all sums over Latin (Greek) indices run from 1 to n (from 1 to N).

The transitions are controlled by different thermal baths at temperatures $T = 1/\beta > 0$ and $T_s = 1/\beta_s > 0$ respectively. Hence the transition probabilities $\rho_{ij|\alpha}$ and $\omega_{\alpha\gamma|i}$ hold the detailed balance conditions (see e.g. [142, 170]):

$$\rho_{ij|\alpha}\; e^{-\beta E_{j\alpha}} = \rho_{ji|\alpha}\; e^{-\beta E_{i\alpha}}, \tag{2.2}$$

$$\omega_{\alpha\gamma|i}\; e^{-\beta_s E_{i\gamma}} = \omega_{\gamma\alpha|i}\; e^{-\beta_s E_{i\alpha}}. \tag{2.3}$$

Without loss of generality we parametrize (2.3) as

$$\omega_{\alpha\,\gamma|i} = e^{B_{\alpha\,\gamma|i} + \frac{\beta_s}{2}(E_{i\,\gamma} - E_{i\,\alpha})}, \qquad B_{\alpha\,\gamma|i} = B_{\gamma\,\alpha|i}, \tag{2.4}$$

where $B_{\alpha\,\gamma|i}$ accounts for the symmetric part of $\alpha \leftrightarrow \gamma$.

2.3 Time-scale separation

Time-scale separation holds when ϵ in (2.1) is sufficiently small; see Figs. (2.1,2.2). This is a reliable approximation, since its predictions are close to the exact stationary probability even for moderately small ϵ; see Fig. (2.1).

We now work out the stationary state ($\dot{p}_{i\alpha} = 0$) of (2.1) for a small ϵ following the standard perturbation theory approach; [192] for a rigorous presentation and [71] for a physical discussion. One puts into (2.1)

$$p_{i\alpha} = B^{[0]}_{i\alpha} + \sum\nolimits_{a\geq 1} \epsilon^a B^{[a]}_{i\alpha}, \tag{2.5}$$

and obtains for successive terms which are of order $\mathcal{O}(1)$ and $\mathcal{O}(\epsilon^a)$, respectively:

$$\sum\nolimits_j [\rho_{ij|\alpha}\, B^{[0]}_{j\alpha} - \rho_{ji|\alpha}\, B^{[0]}_{i\alpha}] = 0, \tag{2.6}$$

$$\sum\nolimits_j [\rho_{ij|\alpha}\, B^{[a]}_{j\alpha} - \rho_{ji|\alpha}\, B^{[a]}_{i\alpha}]$$
$$+ \sum\nolimits_\gamma [\omega_{\alpha\gamma|i}\, B^{[a-1]}_{i\alpha} - \omega_{\gamma\alpha|i}\, B^{[a-1]}_{i\alpha}] = 0, \quad a \geq 1. \tag{2.7}$$

Note that there are solvability conditions found from summing (2.7) over i:

$$\sum\nolimits_{i\gamma} [\omega_{\alpha\gamma|i}\, B^{[a-1]}_{j\alpha} - \omega_{\gamma\alpha|i}\, B^{[a-1]}_{j\alpha}] = 0, \quad a \geq 1. \tag{2.8}$$

Now (2.6) and (2.8) with $a = 1$ are solved as

$$B^{[0]}_{i\alpha} = \bar{p}_{i|\alpha}\bar{p}_\alpha, \tag{2.9}$$

where we used (2.2), the stationary conditional probability $\bar{p}_{i|\alpha}$ of the fast variable has the equilibrium form

$$\bar{p}_{i|\alpha} = e^{-\beta E_{i\alpha}}/Z_\alpha[\beta]\,, \qquad Z_\alpha[\beta] = \sum\nolimits_k e^{-\beta E_{k\alpha}}, \tag{2.10}$$

and where the probability $\bar{p}_\alpha$ of the slow variable is found from (2.8) with $a = 1$

$$\sum\nolimits_\gamma [\bar{\Omega}_{\alpha\gamma}\, p_\gamma - \bar{\Omega}_{\gamma\alpha}\, p_\alpha] = 0, \quad \bar{\Omega}_{\alpha\gamma} \equiv \sum\nolimits_i \omega_{\alpha\gamma|i}\, \bar{p}_{i|\gamma}. \tag{2.11}$$

The meaning of (2.9) is that the fast variable relaxes to the conditional equilibrium, which determines via (2.11) and the effective rates $\bar{\Omega}_{\alpha\gamma}$ the probability of the slow variable.

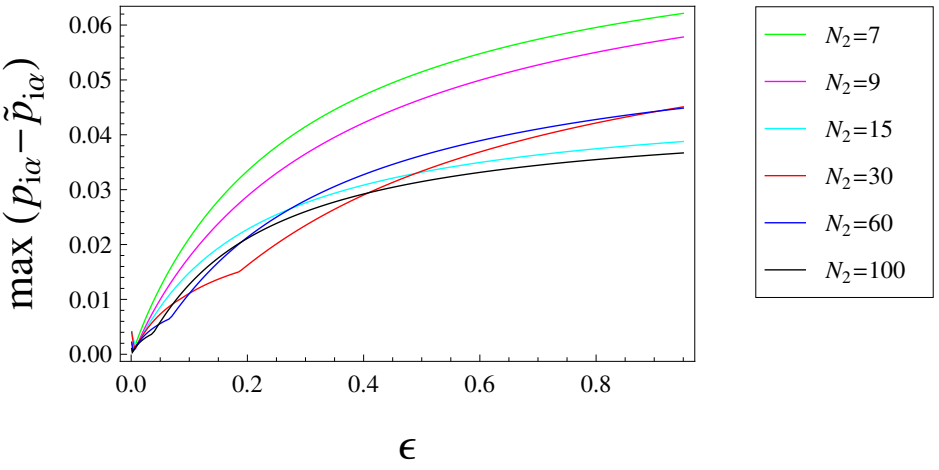

Figure 2.1: $\max(p_{i\alpha} - \widetilde{p}_{i\alpha})$ vs. ϵ, where $p_{i\alpha}$ are the exact stationary probabilities, while $\widetilde{p}_{i\alpha}$ are calculated via the time-scale separation approach. We found that $\max(p_{i\alpha} - \widetilde{p}_{i\alpha}) = \max|p_{i\alpha} - \widetilde{p}_{i\alpha}|$

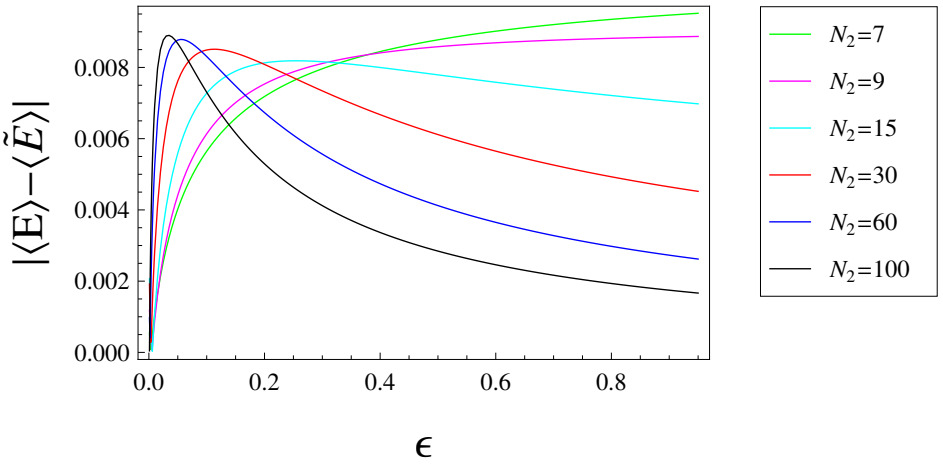

Figure 2.2: $|\langle E\rangle - \langle\widetilde{E}\rangle|$ vs. ϵ, where $\langle E\rangle = \sum_{i\alpha} p_{i\alpha}E_{i\alpha}$ is the exact average energy in the stationary state, $\langle\widetilde{E}\rangle$ is the average energy calculated via the time-scale separation approach, and ϵ controls the slow-fast limit in (2.1). The value of $\langle E\rangle$ (not shown on figure) is some 100–150 times larger than $|\langle E\rangle - \langle\widetilde{E}\rangle|$.

Due to $\sum_{i\alpha} p_{i\alpha} = \sum_{i\alpha} \bar{p}_{i|\alpha}\bar{p}_\alpha = 1$, (2.9) implies normalization condition: $\sum_{i\alpha} B_{i\alpha}^{[a]} = 0$ for $a \geq 1$. Thus $B_{i\alpha}^{[1]}$ is found from this condition, (2.7) with $a = 1$, and (2.8) with $a = 2$. Expectedly, for a small ϵ we get $B_{i\alpha}^{[1]} = \mathcal{O}(\epsilon)$. Now the Fig. (2.1) confirms this fact with numeric results, but it also shows that the difference between (2.9) and the true joint probability is sublinear function of ϵ for a larger values of ϵ. Hence time-scale separation can be applied

beyond the small ϵ situation which is favorable for the approximation

$$p_{i\alpha} = \bar{p}_{i|\alpha}\bar{p}_{\alpha}, \tag{2.12}$$

that we adopt from now on for the stationary probability $p_{i\alpha}$ of (2.1).

In Figs. (2.1,2.2) time-scale separation approach versus numerically exact features of the stationary distribution given by (2.1) is presented. In Eqs. (2.2, 2.3) we took $\beta = 1$, $\beta_s = 10$, $E_{i\alpha} = \alpha i^2/N$ and $n = 2$, where N (n) is the number of different states for the slow (fast) variable. For transition rates in (2.2, 2.3) we choose the Kawasaki rates: $\rho_{ij|\alpha} = e^{\frac{\beta}{2}(E_{j\alpha}-E_{i\alpha})}$ and $\omega_{\alpha\gamma|i} = e^{\frac{\beta_s}{2}(E_{i\gamma}-E_{i\alpha})}$.

2.4 External fields, work and free energies

External fields are introduced via time-dependent parameters $\boldsymbol{a}(t) = (a_1(t), ..., a_A(t))$ in the energy $E_{i\alpha}(\boldsymbol{a})$ of the system. We assume that

$$E_{i\alpha}(\boldsymbol{a}) = E_{\alpha}(\boldsymbol{a}) + \hat{E}_{i\alpha}, \tag{2.13}$$

where $\hat{E}_{i\alpha}$ does not depend on $\boldsymbol{a}$. Eq. (2.13) means that external fields couple only with the slow variable, e.g. because it is difficult to control fast objects. The stationary probabilities (2.12) depend on parameters $\boldsymbol{a}$ through (2.13).

Now $\boldsymbol{a}(t)$ is slow as compared to the relaxation of both fast and slow variables. The temperatures T and T_s are constant. Hence the (quasi-stationary) probabilities of the system are found from (2.12), where $\boldsymbol{a}$ is replaced by $\boldsymbol{a}(t)$. The differential thermodynamic work $\boldsymbol{w}$ is [142] [see (2.13)]

$$\boldsymbol{w} = \sum\nolimits_{i\alpha} \bar{p}_{i\alpha}\, \partial_{\boldsymbol{a}} E_{i\alpha} = \sum\nolimits_{\alpha} \bar{p}_{\alpha}\, \partial_{\boldsymbol{a}} E_{\alpha}, \tag{2.14}$$

where $\partial_{\boldsymbol{a}}$ is the gradient in the $\boldsymbol{a}$-space. Note that relations between thermodynamic and mechanic work have to be specified in the context of concrete applications [193]. This has to

do with the following freedom in the definition (2.14): $\partial_{\boldsymbol{a}} E_\alpha$ (and hence $\boldsymbol{w}$) will change upon adding to E_α a factor $\varphi(\boldsymbol{a})$ that does not depend on α, but depends on $\boldsymbol{a}$ (this is akin to the gauge-freedom of the potential energy in mechanics). Hence the energies E_α need to be specified, before defining the work $\boldsymbol{w}$ [193].

Each component of $\boldsymbol{w} = (w_1, ..., w_A)$ can have a separate physical meaning, since components of $\boldsymbol{a}$ may be driven by different sources.

The integral work is a line integral in the $\boldsymbol{a}$-space:

$$W = \int_{t_{\rm in}}^{t_{\rm f}} {\rm d}t \, \frac{{\rm d}\boldsymbol{a}}{{\rm d}t} \, \boldsymbol{w}(t) = \int_{\boldsymbol{a}_{\rm in}}^{\boldsymbol{a}_{\rm f}} {\rm d}\boldsymbol{a} \, \boldsymbol{w}, \tag{2.15}$$

where $\boldsymbol{a}_{\rm in} = \boldsymbol{a}(t_{\rm in})$ and $\boldsymbol{a}_{\rm f} = \boldsymbol{a}(t_{\rm f})$ are the initial and final values of $\boldsymbol{a}(t)$ reached at times $t_{\rm in}$ and $t_{\rm f}$, respectively. We stress that (2.15) refers to slow changes of parameters $\boldsymbol{a}(t)$, but if we change $\bar{p}_{i\alpha}$ in (2.14) to the time-dependent probability $p_{i\alpha}$ found from (2.1), then the same expression for work applies for arbitrary processes [167].

The work admits a potential (i.e. free energy $\mathbf{F}$) if

$$\boldsymbol{w} = \partial_{\boldsymbol{a}} \mathbf{F}, \qquad W = \mathbf{F}(\boldsymbol{a}_{\rm f}) - \mathbf{F}(\boldsymbol{a}_{\rm in}). \tag{2.16}$$

If (2.16) does not hold, the work extraction (i.e. $W < 0$) by means of a slow, cyclic ($\boldsymbol{a}_{\rm in} = \boldsymbol{a}_{\rm f}$) variation of $\boldsymbol{a}$ is possible. Indeed, if $W \neq 0$, then W changes its sign when the cycle is passed in the opposite direction. The extracted work is determined by the closed-contour integral.

Below we give pertinent examples of $W \neq 0$ for cyclic processes; see (2.41). Now we turn to studying cases, where (2.16) does hold despite of the fact that $T_s \neq T$.

2.5 Non-equilibrium free energy for the activation rate.

An example of slow dynamics is given by the activation energy rate in (2.4) [170]

$$\omega_{\alpha\,\gamma|i} = e^{\beta_s E_{i\gamma}}, \quad B_{\alpha\,\gamma|i} = \beta_s (E_{i\gamma} + E_{i\,\alpha})/2. \tag{2.17}$$

One interpretation of (2.17) is that there is a barrier with energy E^* that is larger than all other energies. Choosing $E^* = 0$, we see that $\omega_{\alpha\,\gamma|i}$ in (2.17) assumes the standard Arrhenius form. The rates (2.17) appears in no-pumping theorem when one considers barrier energies equal to zero [32, 35, 36, 38–41, 43]

Now $\bar{\Omega}_{\alpha\gamma}$ in (2.11) depends only on γ: $\bar{\Omega}_{\alpha\gamma} = \bar{\Omega}_{\gamma}$. The probability $\bar{p}_\alpha$ in (2.11) is found via the detailed balance condition

$$\bar{\Omega}_{\alpha\gamma}\bar{p}_\gamma = \bar{\Omega}_{\gamma\alpha}\bar{p}_\alpha. \tag{2.18}$$

This leads to

$$\bar{p}_\gamma \propto 1/\bar{\Omega}_\gamma. \tag{2.19}$$

Thus we get from (2.10, 2.11, 2.13, 2.17)

$$\bar{p}_\alpha = \mu_\alpha e^{-\beta_s E_\alpha} \Big/ \sum\nolimits_\gamma \mu_\gamma e^{-\beta_s E_\gamma}\,, \tag{2.20}$$

where μ_α corresponds to the weight of the energy E_α and expresses via the statistical sum of the fast variable:

$$\mu_\alpha \;=\; \hat{Z}_\alpha[\beta] \,/\, \hat{Z}_\alpha[\beta - \beta_s], \tag{2.21}$$

$$\hat{Z}_\alpha[\beta] \;\equiv\; \sum\nolimits_k e^{-\beta \hat{E}_{k\alpha}}. \tag{2.22}$$

Note that μ_α does not depend on $\boldsymbol{a}$ due to (2.22) and (2.13). Eqs. (2.13, 2.16, 2.20) imply the existence of a free energy:

$$\mathbf{F} = -T_s \ln\Big[\sum\nolimits_\alpha \mu_\alpha\, e^{-\beta_s E_\alpha}\Big]. \tag{2.23}$$

Recall that $\mathbf{F}$ is defined up to a constant,

$$\mathbf{F} \to \mathbf{F} + \mathcal{C}, \tag{2.24}$$

that can depend on anything besides $\boldsymbol{a}$ [193]. Eq. (2.23) is obtained under a specific choice of $\mathcal{C}$ that proves useful below when calculating derivatives of (2.23).

2.6 Cooling by means of entropy reduction

The free energy $F_{\rm eq}(\boldsymbol{a}) = -T \ln \sum_k e^{-\beta E_k}$ of an equilibrium system is a decreasing function of the temperature:

$$\partial_T F_{\rm eq}(\boldsymbol{a}) = -S_{\rm eq}(\boldsymbol{a}) \leq 0, \tag{2.25}$$

where $S_{\rm eq}(\boldsymbol{a})$ is the entropy. Since $F_{\rm eq}(\boldsymbol{a})$ is defined under a specific choice of an additive *and* temperature-dependent constant [cf. (2.24)], (2.25) changes for a different choice of the constant. Hence (2.25) cannot be interpreted directly. But it can be related to the work-cost of a cooling process via externally-driven parameters $\boldsymbol{a}(t)$ [194].

Normally, cooling means temperature reduction of a macroscopic system that has a single and well-defined temperatures both initially and finally (i.e. an equilibrium system), and then cooling also means that the final entropy (together with the final temperature) is smaller than the initial one. But the needs of NMR-physics [195], atomic and molecular physics [196, 197], quantum computation [198] *etc* led to generalizing this definition [194, 199, 200]. In these fields one needs to reduce the entropy of a system that is coupled to a fixed-temperature thermal bath. Reducing the bath temperature is not feasible. But it is feasible to reduce the entropy via external fields. For instance, in NMR spin systems, entropy decrease for a spin means its polarization increase, which is necessary for the NNR spectroscopy [195].

Thus, cooling amounts to an isothermal process, where $\boldsymbol{a}(t)$ slowly changes from $\boldsymbol{a}_{\rm in}$ to $\boldsymbol{a}_{\rm f}$ and achieves a lower final entropy. For the equilibrium situation this means $S_{\rm eq}(\boldsymbol{a}_{\rm f}) < S_{\rm eq}(\boldsymbol{a}_{\rm in})$ (dynamic aspects of this problem are analyzed in [194]). Now

$$\partial_T [F_{\rm eq}(\boldsymbol{a}_{\rm f}) - F_{\rm eq}(\boldsymbol{a}_{\rm in})] = S_{\rm eq}(\boldsymbol{a}_{\rm in}) - S_{\rm eq}(\boldsymbol{a}_{\rm f}) \geq 0, \tag{2.26}$$

compares two setups at different temperatures, but the same values $\boldsymbol{a}_{\rm in} \to \boldsymbol{a}_{\rm f}$ of the external

fields. Eq. (2.26) means that the work cost $F_{\rm eq}(\boldsymbol{a}_{\rm f}) - F_{\rm eq}(\boldsymbol{a}_{\rm in})$ of cooling increases with the temperature T, i.e. cooling from a higher temperature is harder, as expected.

Turning to the non-equilibrium free energy $\mathbf{F}$, we characterize its temperature dependence via $\partial_{T_s}\mathbf{F}|_T$ and $\partial_T\mathbf{F}|_{T_s}$, since T and T_s are independent parameters. We deduce for the activation energy rate (2.17, 2.21):

$$\begin{aligned}\partial_{T_s}\mathbf{F}|_T &= -S_s - \sum\nolimits_\alpha \bar{p}_\alpha \left[\beta_s \hat{E}_\alpha(\beta-\beta_s) + \ln\mu_\alpha\right] \qquad (2.27)\\ &= -S_s - \sum\nolimits_\alpha \bar{p}_\alpha \int_0^{\beta_s} \mathrm{d}y\,[\,\hat{E}_\alpha(\beta-\beta_s) - \hat{E}_\alpha(\beta-y)\,],\\ \partial_T\mathbf{F}|_{T_s} &= \frac{\beta^2}{\beta_s}\sum\nolimits_\alpha \bar{p}_\alpha \left[\,\hat{E}_\alpha(\beta-\beta_s) - \hat{E}_\alpha(\beta)\right], \qquad (2.28)\end{aligned}$$

where

$$S_s = -\sum_\alpha \bar{p}_\alpha \ln \bar{p}_\alpha \qquad (2.29)$$

is the entropy of the slow variable [cf. (2.25)], and

$$\hat{E}_\alpha = \frac{1}{\hat{Z}(\beta)}\sum_k \hat{E}_{k\alpha} e^{-\beta\hat{E}_{k\alpha}} \qquad (2.30)$$

is the conditionally averaged energy of the fast variable; cf. (2.10). $\hat{E}_\alpha(\beta)$ monotonously decays from $\max_k[E_{k\alpha}]$ to $\min_k[E_{k\alpha}]$ when β goes from $-\infty$ to ∞. Hence we get in (2.27, 2.28):

$$\partial_{T_s}\mathbf{F}|_T \le 0, \qquad \partial_T\mathbf{F}|_{T_s} \ge 0. \qquad (2.31)$$

Now we explore implications of $\partial_T\mathbf{F}|_{T_s} \ge 0$ for a cooling process. Denote by

$$\hat{S}_\alpha(\beta) = -\sum\nolimits_i \bar{p}_{i|\alpha}\ln\bar{p}_{i|\alpha} = \beta\hat{E}_\alpha + \ln\hat{Z}_\alpha, \qquad (2.32)$$

the entropy of the fast variable conditioned by a fixed value α of the slow variable; cf. (2.10). Recall that $S_s + \sum_\alpha \bar{p}_\alpha \hat{S}_\alpha(\beta)$ amounts to the full entropy $-\sum_{i\alpha}\bar{p}_{i\alpha}\ln\bar{p}_{i\alpha}$.

We denote: $\hat{S}_1 = \min_\alpha[\hat{S}_\alpha]$. Then we can define a cooling process, where all energies $E_{\alpha\neq 1}$

in (2.13) slowly increase leading to $\bar{p}_1 \to 1$; see (2.20). Though the process is realized by external fields holding the partial controllability restriction (2.13), we still get a cooling of the *whole* (slow plus fast) system. Indeed, not only the entropy S_s of the slow variable decreases to zero, but also the conditional entropy of the fast variable decreases from its initial value $\sum_\alpha \bar{p}_\alpha \hat{S}_\alpha(\beta)$ to a smaller value $\hat{S}_1(\beta)$.

We note that the considered cooling process can also decrease the internal energy of the system. Recall that the internal energy is defined as [cf. (2.12), (2.13)]

$$\sum\nolimits_{i\alpha} E_{i\alpha} \bar{p}_{i|\alpha} \bar{p}_\alpha = \sum\nolimits_\alpha E_\alpha \bar{p}_\alpha + \sum\nolimits_\alpha \hat{E}_\alpha \bar{p}_\alpha, \tag{2.33}$$

where $\hat{E}_\alpha$ is defined after (2.28). Since $\hat{S}_\alpha$ and $\hat{E}_\alpha$ are equilibrium quantities, $\hat{S}_\alpha$ is an increasing function of $\hat{E}_\alpha$. Hence it is possible to choose $\hat{E}_1 = \min_\alpha[\hat{E}_\alpha]$ in addition to $\hat{S}_1 = \min_\alpha[\hat{S}_\alpha]$. We can also choose $E_1 = \min_\alpha[\hat{E}_\alpha]$. Then $\bar{p}_1 \to 1$ means that the internal energy (2.33) decreases during the cooling.

The cooling process incurs a work cost [cf. (2.23)]

$$\Delta\mathbf{F} = -T_s \ln[\mu_1 e^{-\beta E_1}] - \mathbf{F} \geq 0, \tag{2.34}$$

where $\mathbf{F}$ is the initial free energy. Now $\Delta\mathbf{F} > 0$ means that the work is taken from the external source, hence this is indeed a work cost.

The change of $\Delta\mathbf{F}$ with the temperature T of the fast variable reads from (2.28):

$$\begin{aligned} \partial_T \Delta\mathbf{F}|_{T_s} = -\frac{\beta^2}{\beta_s} \sum\nolimits_\alpha \bar{p}_\alpha \Big[& \hat{E}_\alpha(\beta - \beta_s) - \hat{E}_\alpha(\beta) \\ & - \hat{E}_1(\beta - \beta_s) + \hat{E}_1(\beta) \Big]. \end{aligned} \tag{2.35}$$

Let us now indicate several scenarios for $\partial_T \Delta\mathbf{F}|_{T_s} \leq 0$ in (2.35), and show that they are consistent with condition $\hat{S}_1 = \min_\alpha[\hat{S}_\alpha]$ that defines the cooling process. For example, in (2.35) one can take $\hat{E}_1(\beta - \beta_s) \simeq \hat{E}_1(\beta)$, but $\hat{E}_{\alpha\neq 1}(\beta - \beta_s) \not\simeq \hat{E}_{\alpha\neq 1}(\beta)$. Another example is to make β_s small. Then $E_\alpha(\beta - \beta_s) - \hat{E}_\alpha(\beta) \propto \hat{C}_\alpha$ amounts to the heat-capacity $\hat{C}_\alpha \geq 0$, and then $\partial_T \Delta\mathbf{F}|_{T_s} < 0$ can be implied by $\hat{C}_{\alpha\neq 1} > \hat{C}_1$, which is consistent with $\hat{S}_1 = \min_\alpha[\hat{S}_\alpha]$.

We conclude from $\partial_T \Delta\mathbf{F}|_{T_s} \leq 0$ that it can be easier (in terms of the work-cost) to cool from higher temperatures than from the lower ones. Note that the result survives also in the near-equilibrium limit $\beta \simeq \beta_s$. With the same logics, one shows from (2.27) that when increasing the temperature T_s of the slow variable, one normally has $\partial_{T_s} \Delta\mathbf{F}|_T \geq 0$ [cf. (2.26)], and we revert to equilibrium with $[\partial_{T_s} \Delta\mathbf{F}|_T + \partial_T \Delta\mathbf{F}|_{T_s}]_{T=T_s} \geq 0$.

It is interesting to compare the result of work-cost for entropy reduction with the Mpemba effect [201–203]. The phenomenon was observed independently under various circumstances; apparently, its existence was known already to Aristotle [204]. It can have relatively straightforward physical explanations, e.g. some part of the hotter water can evaporate thus decreasing its amount and making its cooling easier. Another explanation is that a fast cooling can force the water out of local equilibrium [205]. Once such straightforward scenarios are ruled out experimentally, the Mpemba phenomenon disappears [204] (nobody was so far able to replicate the original data by Mpemba and Osborne). It is still possible that the phenomenon survives due to a different chemical composition of a heated water (which may lead to a different concentration of dissolved gases) [206]. At any rate, the phenomenon is not observed once one ensures the identical chemical composition of both samples, e.g. by boiling them first and only then setting them to different initial temperatures [204].

Let us now discuss how our result differs from the above Mpemba phenomenon (without predetermining its causes).

First of all, our cooling is not spontaneous. We focus on an (isothermal) entropy reduction of the system by means of external fields. (Still our cooling process is slow and it leaves the system in its locally stationary state.) Hence we study not the speed of cooling, but its work-cost, as determined by the free-energy. We find that this cost is smaller, if one starts to cool at a higher temperature. Such an effect is impossible in equilibrium, because the equilibrium free energy is a decreasing function of the temperature. However, we show that this effect survives close to equilibrium. Hence the second major difference with the Mpemba phenomenon is that our effect is explicitly out of equilibrium, as it relates to two different temperatures.

2.7 Tree-like topology of the slow variable.

Free energy (2.23) exists for general rates (2.4), if the topology of connections between the states $\alpha, \gamma, ...$ in (2.11) is that of a tree (a network without loops or closed cycles): for a fixed i, $\omega_{\alpha\gamma|i} \neq 0$ (and hence $\bar{\Omega}_{\alpha\gamma} \neq 0$) only along branches of a tree; see Fig. (2.3) for examples. Let node σ_0 be the root of the tree. This root can be chosen arbitrarily, the freedom of choosing

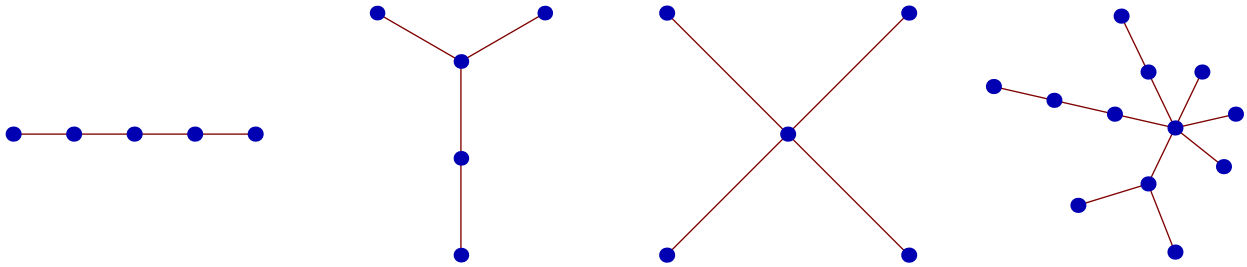

Figure 2.3: Four examples of tree-like structures. Bold points denote states, and lines between them indicate on inter-state transitions.

it will connect to an arbitrary constant (2.24) in the free energy. Each node σ of the tree is related to σ_0 via a unique path $\sigma\sigma'...\sigma''\sigma_0$. The stationary probability $\bar{p}_\sigma$ of σ is deduced from (2.11), it is made of the transition probabilities along this unique path:

$$\bar{p}_\sigma = \frac{\bar{\Omega}_{\sigma\sigma'}\bar{\Omega}_{\sigma'...}\bar{\Omega}_{...\sigma''}\bar{\Omega}_{\sigma''\sigma_0}}{\bar{\Omega}_{\sigma'\sigma}\bar{\Omega}_{...\sigma'}\bar{\Omega}_{\sigma''...}\bar{\Omega}_{\sigma_0\sigma''}}\,\bar{p}_1, \tag{2.36}$$

where $\bar{p}_1$ is gotten from normalization. We add to (2.13):

$$B_{\alpha\,\gamma|i} = B_{\alpha\,\gamma}(\boldsymbol{a}) + \hat{B}_{\alpha\,\gamma|i}, \tag{2.37}$$

where $\hat{B}_{\alpha\,\gamma|i}$ does not depend on $\boldsymbol{a}$. Now $B_{\alpha\,\gamma}(\boldsymbol{a})$ cancels out in (2.36), and (2.36) reduces to (2.20), where [cf. (2.22)]

$$\mu_{\sigma>1} = \frac{\hat{\Omega}_{\sigma\sigma'}\hat{\Omega}_{\sigma'...}\hat{\Omega}_{...\sigma''}\hat{\Omega}_{\sigma''\sigma_0}}{\hat{\Omega}_{\sigma'\sigma}\hat{\Omega}_{...\sigma'}\hat{\Omega}_{\sigma''...}\hat{\Omega}_{\sigma_0\sigma''}}, \qquad \mu_1 = 1, \tag{2.38}$$

$$\hat{\Omega}_{\alpha\gamma} = \frac{1}{\hat{Z}_\gamma[\beta]}\sum\nolimits_i e^{\hat{B}_{\alpha\gamma|i}+\frac{\beta_s}{2}(\hat{E}_{i\gamma}-\hat{E}_{i\alpha})-\beta\hat{E}_{i\gamma}}. \tag{2.39}$$

Hence the free energy (2.23) applies with μ_α given by (2.38). For (2.17), this free energy

differs from (2.23) due to a choice of the constant $\mathcal{C}$ in (2.24).

Note that the tree topology supports the detailed balance: $\bar{\Omega}_{\gamma\alpha}\bar{p}_\alpha = \bar{\Omega}_{\alpha\gamma}\bar{p}_\gamma$. This relates with the Kolmogorov's criterion for the detailed balance: for all loops of the connection network the product of the transition probabilities $\bar{\Omega}_{\gamma\alpha}$ calculated in the clock-wise direction should be equal to the product in the anti-clock-wise direction [54]. For a tree-like network there are no loops, hence the criterion holds.

For general rates already one loop (i.e. a three-level system) may suffice for invalidating the existence free energy. But we expect that loops will be less relevant for multi-dimensional models. Indeed, let us return to $\bar{p}_\alpha \propto \mu_\alpha e^{-\beta E_\alpha}$, $\bar{p}_{i\alpha} \propto e^{-\beta \hat{E}_{i\alpha}}$, but instead of (2.13) we assume that only one externally-driven parameter (out of two) pertains to the slow variable:

$$E_{i\alpha}(\boldsymbol{a}) = E_\alpha(a_1) + \hat{E}_{i\alpha}(a_2). \tag{2.40}$$

Eq. (2.40) is the minimal situation, where the work-extraction via a slow, cyclic process is possible.

We get for the rotor of the work $\boldsymbol{w}$ [cf. (2.14) and (2.22)]:

$$\begin{aligned} \partial_{a_2} w_{a_1} - \partial_{a_1} w_{a_2} = &\langle\, (\partial_{a_1} E\,)\, \partial_{a_2} [\,\ln \mu - T\beta_s \ln \hat{Z}\,]\,\rangle \\ &-\langle\, \partial_{a_1} E\rangle\, \langle\, \partial_{a_2} [\,\ln \mu - T\beta_s \ln \hat{Z}\,]\,\rangle, \end{aligned} \tag{2.41}$$

where $\langle X\rangle \equiv \sum_\alpha \bar{p}_\alpha X_\alpha$. Eq. (2.41) shows how far is the work from having a gradient when the partial controllability (2.13) does not hold. It also determines the amount of work extracted from a cycle in the (a_1, a_2)-space. Now since (2.41) is a correlation, the existence of a gradient for $w_{\boldsymbol{a}}$ can be recovered if fluctuations are negligible.

2.8 Quasi-continuous limit for the slow variable.

Recall that the index $\gamma = 1, ..., N$ numbers the states of the slow variable. We now consider the case, where these states are arranged over a segment of a line (1D situation), so that only transition from one neighbour state to another are allowed. This is an example of

birth-death processes that have many applications [9]. If the segment is finite and the states are homogeneously and densely located in it, then one can pass to the continuous limit, where instead of a discrete index γ we shall have a continuous variable x [22, 156]. The continuous limit is achieved by

$$E_{i\,\gamma} \to E_i(x), \quad E_{i\,\gamma+1} \to E_i(x) + \hat{\epsilon} E_i'(x), \tag{2.42}$$

where x is a continuous parameter, $\hat{\epsilon}$ is a small parameter (the distance between the states γ and $\gamma+1$ on the segment), and where $A'(x) \equiv \mathrm{d}A(x)/\mathrm{d}x$. We also denote

$$B_{\gamma+1\,\gamma|i} \to B_i(x). \tag{2.43}$$

Since $E_{i\,\gamma+1} \approx E_{i\,\gamma}$ we expand in (2.4) over a small $\hat{\epsilon}$

$$\omega_{\gamma+1\,\gamma|i} = \exp\left[B_{\gamma+1\,\gamma|i}\right](1 + \beta_s(E_{i\,\gamma} - E_{i\,\gamma+1})/2), \tag{2.44}$$

and get in (2.1) (see [22, 156] for similar derivations):

$$\begin{aligned}
&\sum\nolimits_{\gamma}[\omega_{\alpha\gamma|i}\, p_{i\gamma} - \omega_{\gamma\alpha|i}\, p_{i\alpha}] \\
&= \beta_s \left[e^{B_{\alpha\,\alpha+1|i}} (E_{i\,\alpha+1} - E_{i\,\alpha}) \frac{p_{i\,\alpha+1} + p_{i\,\alpha}}{2} - e^{B_{\alpha-1\,\alpha|i}} (E_{i\,\alpha} - E_{i\,\alpha-1}) \frac{p_{i\,\alpha-1} + p_{i\,\alpha}}{2} \right] \\
&+ e^{B_{\alpha+1\,\alpha|i}} (p_{i\,\alpha+1} - p_{i\,\alpha}) - e^{B_{\alpha\,\alpha-1|i}} (p_{i\,\alpha} - p_{i\,\alpha-1}) \\
&= \beta_s [\zeta_{i\,\alpha+1} - \zeta_{i\,\alpha}] + \xi_{i\,\alpha+1} - \xi_{i\,\alpha},
\end{aligned} \tag{2.45}$$

where we denoted

$$\begin{aligned}
\xi_{i\,\alpha+1} &\equiv e^{B_{\alpha+1\,\alpha|i}} (p_{i\,\alpha+1} - p_{i\,\alpha}), \\
\zeta_{i\,\alpha+1} &\equiv e^{B_{\alpha\,\alpha+1|i}} (E_{i\,\alpha+1} - E_{i\,\alpha})\, (p_{i\,\alpha} + p_{i\,\alpha+1})/2.
\end{aligned} \tag{2.46}$$

Hence (2.45) can be written [see(2.42,2.43)]

$$\beta_s[\zeta_{i\,\alpha+1} - \zeta_{i\,\alpha}] + \xi_{i\,\alpha+1} - \xi_{i\,\alpha} = \hat{\epsilon}[\beta_s \partial_x \zeta_i(x) + \partial_x \xi_i(x)] \tag{2.47}$$

Note that $\zeta_i(x)$ and $\xi_i(x)$ are proportional to $\hat{\epsilon}$

$$\begin{aligned} \zeta_i(x) &= \hat{\epsilon} e^{B_i(x)} p_i(x) \partial_x E_i(x), \\ \xi_i(x) &= \hat{\epsilon} e^{B_i(x)} \partial_x p_i(x) \end{aligned} \tag{2.48}$$

Eqs.(2.1, 2.45, 2.47, 2.48) lead to:

$$\dot{p}_i(x,t) = \sum\nolimits_j [\rho_{ij|x}\, p_j(x) - \rho_{ji|x}\, p_i(x,t)] + \epsilon\, \hat{\epsilon}^2 \partial_x J_i(x,t), \tag{2.49}$$

$$J_i(x,t) \equiv \beta_s\, e^{B_i(x)}\, p_i(x,t)\, E_i'(x) + e^{B_i(x)}\, p_i'(x,t), \tag{2.50}$$

where $p_i(x)$ is the joint probability of i and x: $\sum_i \int \mathrm{d}x\, p_i(x,t) = 1$, and $J_i(x,t)$ is the probability current related to the slow variable.

Eq. (2.49) shows that as compared to (2.1) the slow-fast limit is facilitated due to the additional small factor $\hat{\epsilon}^2$. This is confirmed by Fig. 1 which shows that the slow-fast limit improves when the number N of states of the slow variable is large. Thus the above continuous limit is a way to get the time-scale separation naturally.

We treat (2.49) wih the same time-scale separation argument (2.5–2.11) with minor modifications for the continuous case. The stationary conditional probability is still $\bar{p}_{i|x} \propto e^{-\beta E_i(x)}$. For the stationary probability density $\bar{p}(x)$|which holds the zero-current condition $\sum_i J_i(x) = 0$|we obtain from (2.49, 2.50):

$$\begin{aligned} &\bar{p}(x) \propto e^{-\beta_s E(x)} \mu(x), \qquad \mu(x) = e^{h(x)} \sum\nolimits_i e^{-\beta \hat{E}_i(x)}, \\ &h'(x) = (\beta - \beta_s) \frac{\sum_i \hat{E}_i'(x) e^{\hat{B}_i(x) - \beta \hat{E}_i(x)}}{\sum_i e^{\hat{B}_i(x) - \beta \hat{E}_i(x)}}, \end{aligned} \tag{2.51}$$

where by analogy to (2.13, 2.37): $E_i(x;\boldsymbol{a}) = E(x;\boldsymbol{a}) + \hat{E}_i(x)$, $B_i(x;\boldsymbol{a}) = B(x;\boldsymbol{a}) + \hat{B}_i(x)$. The free energy is given as in (2.23) by $\mathbf{F} = -T_s \ln \int \mathrm{d}x\, \mu(x)\, e^{-\beta_s E(x)}$. Note that it goes to

the equilibrium expression for $T_s = T$.

A particular case of (2.51) is when $B_i(x)$ does not depend on i, i.e. $\hat{B}_i(x) = 0$. Now the non-equilibrium free energy reads [183]:

$$\mathbf{F}^\circ = -T_s \ln \left[\int \mathrm{d}x \, \big(\sum\nolimits_i e^{-\beta E_i(x)} \big)^{\beta_s/\beta} \right]. \tag{2.52}$$

The following features of $\mathbf{F}^\circ$ are deduced directly from (2.52) [183]. They all are very similar to those of the equilibrium free energy.

– $\mathbf{F}^\circ$ is the potential for the work *without* restriction (2.13), i.e. now $\boldsymbol{a}$ can also enter $\hat{E}_i(x)$:

$$E_i(x;\boldsymbol{a}) = E(x;\boldsymbol{a}) + \hat{E}_i(x;\boldsymbol{a}). \tag{2.53}$$

– For the temperature-derivatives of we get

$$\partial_T \mathbf{F}^\circ|_{T_s} = -S, \qquad \partial_{T_s} \mathbf{F}^\circ|_T = -S_s, \tag{2.54}$$

where $S \equiv -\int \mathrm{d}x \sum_k \bar{p}_k(x) \ln \bar{p}_{k|x}$ and $S_s \equiv -\int \mathrm{d}x \, \bar{p}(x) \ln \bar{p}(x)$ are, respectively, the conditional entropy of the fast variable, and the marginal entropy of the slow variable. Hence the anti-thermodynamic cooling effect noted in (2.35) is impossible here.

Note the analogy between (2.54) and the equilibrium formula $\partial_T F_{\mathrm{eq}} = -S_{\mathrm{eq}}$. It leads us to $\mathbf{F}^\circ = U - T_s S_s - TS$, where $U = \int \mathrm{d}x \sum_k \bar{p}_k(x) E_k(x)$ is the average (overall) energy. This expression for $\mathbf{F}^\circ$ already appeared in the physics of Brownian motion [182, 183] and glasses [175].

Eq. (2.52) has an intuitively appealing meaning [182–189], since it implies that the free energy $-T \sum_i e^{-\beta E_i(x)}$ of the fast variable (evaluated at a fixed value of the slow variable) serves as an effective potential $U_{\mathrm{eff}}(x)$ for the (Gibbs) distribution of the slow variable: $\bar{p}(x) \propto e^{-\beta_s U_{\mathrm{eff}}(x)}$. And then $\mathbf{F}^\circ$ is the free energy related to that effective Gibbs distribution: $\mathbf{F}^\circ = -T_s \ln \int \mathrm{d}x \, e^{-\beta_s U_{\mathrm{eff}}(x)}$. While this heuristic explanation is frequently applied in statistical physics, we should keep in mind from the above derivation that $\mathbf{F}^\circ$ exists *due* to the continuous

limit, e.g. assuming $\hat{B}_{\alpha\gamma} = 0$ in the discrete case does not recover the above equilibrium features (for $T \neq T_s$).

Chapter 3

Adaptive heat engine [1]

3.1 Introduction

Heat-engines drove the Industrial Revolution and their foundation, *viz.* thermodynamics, became one of the most successful physical theories [207]. Extensions of thermodynamics to stochastic [58, 170] and quantum domain [86] led to new generations of heat engines [58, 86, 88–100]. As everyone could observe, the work-extraction function of macroscopic heat-engines requires external on-line control, e.g. the Carnot cycle performing the optimal-efficiency cycle do demand a careful external control of its functioning, one should provide the specific sequence of adiabatic and isothermal processes [207, 208]. Smaller engines studied within stochastic or quantum thermodynamics may not demand on-line control, i.e. they are autonomous [96, 97], but they do demand fitting between internal and environmental parameters [89, 90, 92–95, 197], e.g. because for fixed environment (thermal baths) there are internal parameters, under which the machine acts as a heat-pump or refrigerator performing tasks just opposite to that of heat-engine. Such fitted engines are susceptible to environmental changes, e.g. when the bath temperatures get closer due to the very engine functioning. Car engines treat this problem by abandoning the partially depleted fuel (i.e. the hot bath), and using fresh fuel.

[1] The results considered in this chapter are published in Ref. [120].

Here we study a rudimentary model of autonomous, adaptive heat engine. Adaptive means that the engine can work for a sufficiently general class of environments, i.e. it needs neither on-line control, nor an externally imposed fitting between its internal parameters and the bath temperatures. In particular, the engine can adapt to the results of its own functioning. Hence adaptive engines can be useful for fueling devices via unknown or scarce resources [209].

The major biophysical heat engine, *viz* photosynthesis|which operates between the hot Sun temperature and the low-temperature Earth environment [210]||does have adaptive features that allow its functioning under decreased hot temperature (shadowing) or increased cold temperature (hot whether) [121, 122].

Recently, several physical models concentrated on adaptive sensors, adaptive transport models *etc* [122, 176, 211–216]. These studies clarified thermodynamic costs of adaptation scenarios [122, 176, 213–216]. Other research lines related adaptation with (poly)homeostasis [217] and models of artificial life [208, 218].

For analyzing the adaptation and its resources for heat engines, we need a tractable and realistic model that is much simpler than e.g. its prototypes in photosynthesis. The model ought to consist of the proper heat-engine and a controller that ensures the adaptation; see Fig. 3.1. Together they form an authonomous system.

3.2 The functional degree of freedom

We choose one of the most known models of quantum/stochastic thermodynamics that was introduced and studied as a model for maser [89, 90, 92, 93, 197]. Related models were studied in the context of photovoltaics [98]. The model has three states: $i = 1, 2, 3$. Each state i has energy E_i. Transitions between different states are caused by thermal baths that can provide or accept necessary energies. We assume that the resulting dynamics is described by a Markov master equation [9]:

$$\dot{p}_i \equiv \mathrm{d}p_i/\mathrm{d}t = \sum\nolimits_j [\rho_{i\leftarrow j} p_j - \rho_{j\leftarrow i} p_i], \quad i, j = 1, 2, 3, \tag{3.1}$$

where $p_i(t)$ is the probability of the state i at time t, and $\rho_{i\leftarrow j} > 0$ is the transition rate from j to i. We assume for simplicity that each transition $i \leftrightarrow j$ couples with one equilibrium bath at temperature $T_{ij} = T_{ji} = 1/\beta_{ji}$; see Fig. 3.1. The equilibrium nature of each bath imposes the detailed balance constraint for transitions [9]:

$$\rho_{i\leftarrow j}\, e^{-\beta_{ij}E_j} = \rho_{j\leftarrow i}\, e^{-\beta_{ij}E_i}, \qquad \beta_{ij} = \beta_{ji}. \tag{3.2}$$

We take one temperature infinite: $\beta_{21} = 0$. This bath is then a work-source, because due to $\mathrm{d}S_{21} = \beta_{21}\mathrm{d}Q_{21} = 0$ it exchanges energy $\mathrm{d}Q_{21} \neq 0$ at zero entropy change $\mathrm{d}S_{21} = 0$. The other two thermal baths are the ones necessary for any heat-egine; see Fig. 3.1. Below we ensure its function for a large range of $\beta_{31} \neq \beta_{32}$.

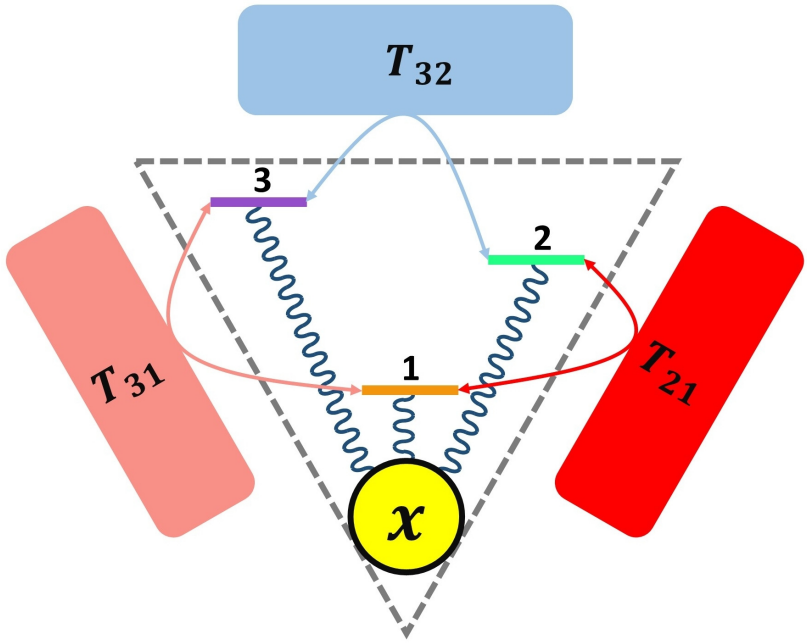

Figure 3.1: A schematic representation of the model. There are three thermal baths at temperatures $T_{32}, T_{31} < T_{21} = \infty$; each one drives a single transition among three engine levels 1, 2, and 3. The bath with temperature $T_{21} = \infty$ is the source of work. A controller x interacts with energies, but does not couple directly with the baths.

Since each bath causes only one transition $i \leftrightarrow j$,

$$K_{ij} = K_{ji} = (E_i - E_j)(\rho_{i\leftarrow j}p_j - \rho_{j\leftarrow i}p_i), \tag{3.3}$$

is the average energy lost (if $K_{ij} > 0$) or gained (if $K_{ij} < 0$) by the bath per time-unit; see

(3.1). In the stationary state these energy currents K_{ij} hold $K_{31}+K_{32}+K_{21} = \sum_{i=1}^{3} \dot{p}_i E_i = 0$, as necessary for the average energy conservation. Eq. (3.1) implies for stationary probabilities

$$\begin{aligned} p_1 &= \frac{1}{\mathcal{Z}}[\rho_{1\leftarrow 2}\rho_{1\leftarrow 3} + \rho_{3\leftarrow 2}\rho_{1\leftarrow 3} + \rho_{2\leftarrow 3}\rho_{1\leftarrow 2}], \\ p_2 &= \frac{1}{\mathcal{Z}}[\rho_{2\leftarrow 1}\rho_{2\leftarrow 3} + \rho_{3\leftarrow 1}\rho_{2\leftarrow 3} + \rho_{2\leftarrow 1}\rho_{1\leftarrow 3}], \\ p_3 &= \frac{1}{\mathcal{Z}}[\rho_{3\leftarrow 1}\rho_{3\leftarrow 2} + \rho_{2\leftarrow 1}\rho_{3\leftarrow 2} + \rho_{1\leftarrow 2}\rho_{3\leftarrow 1}] \end{aligned} \tag{3.4}$$

where $\mathcal{Z}$ ensures $\sum_{i=1}^{3} p_i = 1$. Using (3.2–3.2) and $\rho_{1\leftarrow 2} = \rho_{2\leftarrow 1}$ due to $\beta_{21} = 0$ we get

$$K_{21} = \frac{\hat{E}_2}{\mathcal{Z}}\,\rho_{2\leftarrow 1}\,\rho_{1\leftarrow 3}\,\rho_{3\leftarrow 2}\left[1 - e^{(\beta_{32}-\beta_{31})\hat{E}_3 - \beta_{32}\hat{E}_2}\right], \tag{3.5}$$

$$K_{31} = -\hat{E}_3 K_{21}/\hat{E}_2, \qquad K_{32} = (\hat{E}_3 - \hat{E}_2)K_{21}/\hat{E}_2, \tag{3.6}$$

$$\hat{E}_2 \equiv E_2 - E_1, \qquad \hat{E}_3 \equiv E_3 - E_1. \tag{3.7}$$

The heat-engine functioning is defined as [cf. (3.3)]

$$0 > K_{21} = -(E_2 - E_1)(p_2 - p_1)\rho_{1\leftarrow 2}, \tag{3.8}$$

i.e. the energy goes to the work-source with the *power* $|K_{21}|$. Inequality (3.8) shows that the heat-engine functions via population inversion between energy levels E_1 and E_2: when driving the transition $1 \leftrightarrow 2$, the work-source gains energy in average. Using (3.5, 3.7) we write (3.8) as

$$\hat{E}_2[(1-\theta)\hat{E}_3 - \hat{E}_2] > 0, \quad \theta \equiv \beta_{31}/\beta_{32}. \tag{3.9}$$

Eq. (3.9) demands different temperatures: $\beta_{32} \neq \beta_{31}$. It also demands tuning between the energies $\hat{E}_2$, $\hat{E}_3$ and θ: it is impossible to hold (3.9) for a wide range of θ by means of constant $\hat{E}_2$ and $\hat{E}_3$; e.g. if (3.9) holds for $1 > \theta$ due to $\hat{E}_3 > \hat{E}_2 > 0$, then it is violated for $1 - \theta < \frac{\hat{E}_2}{\hat{E}_3}$. Tuning is necessary, since for suitable values of $\hat{E}_2$ and $\hat{E}_3$, the machine can function also as a refrigerator or as a heat-pump.

The efficiency (power divided over the incoming current) η of the engine is given as (see

(3.5, 3.6, 3.9))

$$\eta \equiv \frac{-K_{21}}{\max[K_{31}, K_{32}]} \leq \eta_{\rm C} \equiv 1 - \min[\theta, \frac{1}{\theta}], \tag{3.10}$$

$$\eta = \max\left[\frac{\hat{E}_2}{\hat{E}_3}, \frac{\hat{E}_2}{\hat{E}_2 - \hat{E}_3}\right] \tag{3.11}$$

where η depends only on energy differences $\hat{E}_2$ and $\hat{E}_3$ [see (3.6)], and the Carnot value $\eta_{\rm C}$ bounds η from above, as deduced from (3.9). For $\eta \to \eta_{\rm C}$ inequality (3.9) saturates, and (3.5, 3.6) show that all K_{ij} nullify (power-efficiency trade-off) [58, 86, 95]; see [219] for most recent discussion of this trade-off.

3.3 The structural degree of freedom

The controller x should ensure adaptation to continuous environmental variations; hence it is continuous. x and i interact via energies $E_i(x)$. The joint probability $p_i(x,t)$ of x and i, $\int {\rm d}x \sum_i p_i(x,t) = 1$, evolves via the Fokker-Planck plus master equations [cf. (3.1)] [9]:

$$\begin{aligned}\dot{p}_i(x,t) = &\sum\nolimits_{j=1}^{3}[\rho_{i\leftarrow j}(x)p_j(x,t) - \rho_{j\leftarrow i}(x)p_i(x,t)] \\ &+\frac{1}{\gamma}\partial_x[p_i(x,t)E_i'(x)] + D\partial_x^2 p_i(x,t), \quad i = 1,2,3,\end{aligned} \tag{3.12}$$

where $E_i'(x) \equiv \frac{{\rm d}E_i(x)}{{\rm d}x}$, $\gamma > 0$ is the friction constant, and $D > 0$ is the diffusion constant. $\rho_{i\leftarrow j}(x)$ is specified in (3.22); it holds (3.2) with $E_i \to E_i(x)$ and $E_j \to E_j(x)$.

Eq. (3.12) has a wide range of chemical and biological applications [156, 220–222, 222–229]. It accounts for enzyme dynamics, where a reaction (to be accelerated) is described by a discrete variable i that interacts with a coarse-grained conformational coordinate x of the enzyme [222–228]. The existence of x was inquired from experiments [156, 220–222, 222–226], and also deduced from microscopic models [227, 228]. Reactions of photosynthesis are also known to interact with conformational degrees of freedom [230]: the main mechanism of the photosynthesis adaptation was located in conformational changes of the thylakoid membrane that bounds light-dependent reactions [121].

Eq. (3.12) has similarities with recent models for quantum heat engines; but there the continuous variable is employed for storing the extracted work [99, 100].

The harmonic choice of interaction energies $E_i(x)$ verified itself well in various applications [222, 223, 226, 227, 229]

$$E_i(x) = a(x - b_i)^2 + c_i, \quad i = 1, 2, 3, \tag{3.13}$$

where $a > 0$, b_i and c_i are constants; a is i-independent, since $E_i(x)$ have the same shape for $x \to \pm\infty$. Eq. (3.13) was explained in [228] via quantum chemistry.

The exact stationary solution $p_i(x)$ of (3.12) is not found. We need an explicit form of $p_i(x)$, since from within $p_i(x)$ we should search for adaptation-friendly shapes of $E_i(x)$. This issue is solved, if we assume in (3.12) that x is slow: $\frac{1}{\gamma}$, $D \ll \rho_{i\leftarrow j}(x)$. This is realistic for enzymes, where x includes large molecular groups whose motion is slow [225, 226]. A virtue of the slow limit is that it decreases energy costs related to control, akin to the standard reversibility limit of thermodynamics; see below. The slow limit can be justified introducing in (3.12) the conditional probability $p_{i|x}(t)$ [235],

$$p_i(x,t) = p_{i|x}(t)p(x,t), \quad \int \mathrm{d}x\, p(x,t) = 1, \quad \sum_{i=1}^{3} p_{i|x}(t) = 1,$$

and collecting fast terms:

$$\dot{p}_{i|x} = \sum\nolimits_{j=1}^{3} [\rho_{i\leftarrow j}(x)p_{j|x} - \rho_{j\leftarrow i}(x)p_{i|x}]. \tag{3.14}$$

Slow terms are found from (3.12, 3.14) by summing over i:

$$\dot{p}(x,t) = \frac{1}{\gamma}\partial_x[p(x,t)\sum_{i=1}^{3} p_{i|x}E_i'(x)] + D\partial_x^2 p(x,t). \tag{3.15}$$

Since i is fast, $p_{i|x}$ in (3.15) can be taken as time-independent, i.e. $p_{i|x}$ is found from (3.2) upon replacing there $\rho_{ij} \to \rho_{ij}(x)$ [235]. The stationary probability of x is found from (3.15)

via the zero-current condition $p(x)\sum_{i=1}^{3} p_{i|x} E_i'(x) + \gamma D \partial_x p(x) = 0$:

$$p(x) \propto e^{-\Psi(x)/(\gamma D)}, \quad \Psi'(x) \equiv \sum\nolimits_{i=1}^{3} p_{i|x} E_i'(x), \tag{3.16}$$

where $\Psi'(x) = \frac{\mathrm{d}\Psi}{\mathrm{d}x}$. Using $\sum_{i=1}^{3} p_{i|x} = 1$ we define [cf. (3.7)]

$$\Phi'(x) \equiv \sum\nolimits_{i=2}^{3} p_{i|x} \hat{E}_i'(x) = \Psi'(x) - E_1'(x), \tag{3.17}$$

$$\hat{E}_i(x) \equiv E_i(x) - E_1(x). \tag{3.18}$$

3.4 Adaptation

The energies $E_i(x)$ do not depend on β_{31} and β_{32}. We choose $E_i(x)$ as follows. First, there is a unique maximally probable value $\hat{x}$ of x, which is the minimum of the effective potential $\Psi(x)$, i.e. [from (3.16, 3.17)]:

$$\Phi'(\hat{x}) = -E_1'(\hat{x}), \qquad \Phi''(\hat{x}) > -E_1''(\hat{x}), \tag{3.19}$$

where the latter condition means stability.

Second, the heat-engine condition $K_{21}(x) < 0$ holds in a vicinity of the maximally probable value $\hat{x}$ [cf. (3.9, 3.18)]:

$$\hat{E}_2(x)[(1-\theta)\hat{E}_3(x) - \hat{E}_2(x)] > 0, \quad x \simeq \hat{x}, \tag{3.20}$$

Work-extraction should also hold in average [cf. (3.2, 3.3)]:

$$\langle K_{21} \rangle = \rho_{1\leftarrow 2} \int \mathrm{d}x \, \hat{E}_2(x) p(x) (p_{1|x} - p_{2|x}) < 0. \tag{3.21}$$

Eq. (3.20) implies (3.21), if γD is small enough, because in this limit the probability of x concentrates around $\hat{x}$; see (3.16). More generally (e.g. out of the slow limit), the adaptation criterion can be based directly on (3.21).

The adaptation scenario implied by (3.17, 3.20) is as follows. Assume that the machine in its stationary state is already working as a heat-engine for certain bath temperatures β_{31} and β_{32}. If one of them (or both) change, $p_{i|x}$ goes out of the stationary state. As shown by (3.17), this moves $\hat{x}$ to a new value, where according to (3.20) the heat-engine function is recovered. We stress that in this feedback scheme, x does not interact directly with the baths: the change of $\hat{x}$ comes from $p_{i|x}$; see (3.17). Using a feed-forward scheme, where x directly couples with the baths, does not lead to advantages with respect to adaptation, because it changes only friction and diffusion in (3.12). It is possible to study scenarios, where in addition to the heat-engine function (3.20) the adaptation optimizes the heat-engine efficiency η or its power $|K_{21}|$.

To search for adaptation, we focus on the following class of transition rates [cf. (3.1, 3.2)]:

$$\rho_{i\leftarrow j}(x) = f_{ij}[\, E_j(x) - E_i(x)\,], \tag{3.22}$$

where $f_{ij}[y]$ holds (3.2). Eq. (3.22) includes the Kramers' rate $f_{ij}[y] = e^{\beta_{ij}[\Delta_{ij}+\min(y,0)]}$, where $\Delta_{ij} = \Delta_{ji}$ is the barrier height [9], and $f_{ij}[y] = e^{\beta_{ij}y/2}$ that relates to the discrete-space Fokker-Planck equation [156]. The constraint (3.22) relates to one of conditions of the no-pumping theorem [32, 35, 36, 38–41, 43].

Eqs. (3.22, 3.18, 3.2) imply that $\rho_{ij}(x)$, the stationary $p_{i|x}$, and $\Phi'(x)$ in (3.17) depend on $E_1(x)$ only via $\hat{E}_3(x)$ and $\hat{E}_2(x)$. Hence we study $\Phi'(x)$ for given $\hat{E}_3(x)$ and $\hat{E}_2(x)$, look for a domain where (3.20) holds, and then define $\hat{x}$ via (3.19) by choosing a suitable $E_1(x)$ that does not depend on β_{31} and on β_{32}. This is achieved by plotting $\Phi'(x)$ as a function of x under different values of β_{31} and β_{32}.

Note that (3.20) confines the shape of $\hat{E}_2(x)$: since (3.20) should hold for $\theta \to 1$, there exists x_0 such that $\hat{x} \to x_0$ for $\theta \to 1$, and $\hat{E}_2(x_0) = 0$. In the vicinity of x_0, $\hat{E}_3(x)$ is either finite or goes to zero slower than $\hat{E}_2(x)$, so that (3.20) still holds for $\theta \to 1$ and $\hat{x} \to x_0$.

3.5 Restricted adaptation

Let us first assume that one temperature (say β_{31}) takes arbitrary positive values, while another one (β_{32}) is fixed. Adaptation is necessary here, since $\theta = \beta_{31}/\beta_{32}$ is an arbtrary positive number, hence (3.20) cannot be valid for x-independent E_i. Applying the above method, we deduce that (3.17, 3.20) for adaptation can be satisfied for the experimentally motivated choice (3.13) for $E_i(x)$; see Fig. 3.2.

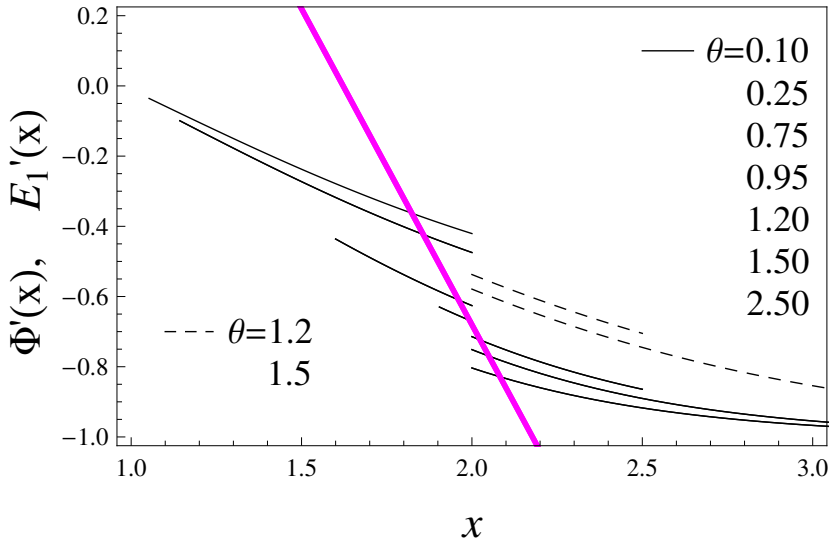

Figure 3.2: Restricted adaptation scenario. $\Phi'(x)$ given by (3.17, 3.22) with $f_{ij}[y] = e^{\beta_{ij}y/2}$. (Similar results hold for all other physical choices of $f_{ij}[y]$; see (3.22).) We assume $\hat{E}_3(x) = -x$, $\hat{E}_2(x) = x - 2$; see (3.18, 3.13). Now heat-engine conditions (3.20) hold for $x > 2$ if $\theta > 2$, and for $x \in (\frac{2}{2-\theta}, 2)$ if $\theta < 2$. Normal (resp. dashed) curves: $\Phi'(x)$ for x that support (3.20) under $\beta_{32} = 1$ (resp. $\beta_{32} = 0.7$) and various $\theta = \beta_{31}/\beta_{32}$. They are indicated from the top to the bottom in the right (resp. left). The magenta (bold) curve shows $-E_1'(x)$, where $E_1'(x) = 1.8(x-2) + 0.680289$; cf. (3.13). Intersections of $-E_1'(x)$ with $\Phi'(x)$ determine $\hat{x}$. Conditions (3.19) hold for all normal curves, and none of dashed curves.

Since the validity domain (3.20) of the heat-engine shrinks to a point for $\theta \to 1$, we need progressively smaller values of $D\gamma$ in (3.16) for ensuring the average work-extraction (3.21) under $\theta \to 1$. If the diffusion of x is caused by an equilibrium bath, we get $D\gamma = T$ in (3.12) [9], and the temperature T of this bath should be sufficiently low for (3.21) to hold. If this is the lowest temperature, there is a heat current $K_{\rm A}$ towards it tending to increase it. Hence this low temperature is a resource; see [176, 213–216] for related results. In the slow limit, $K_{\rm A} = \mathcal{O}(\frac{1}{\gamma})$ can be much smaller than the energy currents K_{ij} of the heat-egine. Eqs. (3.6, 3.9 3.10) imply that $K_{ij} \to 0$ for high efficiencies $\eta \to \eta_{\rm C}$. In that case $K_{\rm A}$ stays

finite and is the dominant energy current.

3.6 Full adaptation

Let now *both* β_{31} and β_{32} vary. Fig. 3.2 shows that the set-up which worked for a fixed β_{32} does not apply: adaptation conditions (3.19) break down in a vicinity of $\theta = 1$. Topologically, changing both β_{32} and β_{31} destroys fine-tuned equality in (3.19); see Fig. 3.2.

We argue that adaptation conditions (3.19) cannot hold together, if both β_{31} and β_{32} vary, *and* if $\frac{\mathrm{d}}{\mathrm{d}y} f_{ij}[y] \geq 0$; see (3.22). The latter holds for all physical examples we are aware of, and means that the transition from one energy to another is facilitated, if the lower energy increases or the higher energy decreases.

The only way we found for recovering adaptation is to assume that x is subject to a *negative friction*: $\gamma < 0$. Note that the $\gamma < 0$ and $D > 0$ situation is stable, since (3.15) does predict relaxation to (3.16).

One way to achieve $\gamma < 0$ is to subject x to a negative-temperature (population-inverted) thermal bath: $\beta < 0$. Eqs. (3.12) with $\gamma \propto \beta < 0$ is then an effective description of quasi-continuous, but discrete degrees of freedom. Negative temperatures are known for various systems whose energies are bounded from above [236–240]. Now $\beta < 0$ is a resource, since when coupled to positive temperatures, β tends to increase [236]. Physically, $\beta < 0$ means a reservoir of stored energy, and the fact that β decreases means that this energy is spent for adaptation.

Other examples of negative friction include negative resistance of electric circuits [232], negative viscosity of driven fluids [233], and the negative absolute mobility for Brownian systems [234].

Now $D\gamma < 0$ in (3.16), and the most probable $\hat{x}$ means that inequalities in (3.19) are reversed. New adaptation conditions are (3.20) and [instead of (3.19)]

$$\Phi'(\hat{x}) = -E_1'(\hat{x}), \qquad \Phi''(\hat{x}) < -E_1''(\hat{x}). \tag{3.23}$$

These conditions can be satisfied, as seen in Fig. 3.3. In contrast to the previous scenario,

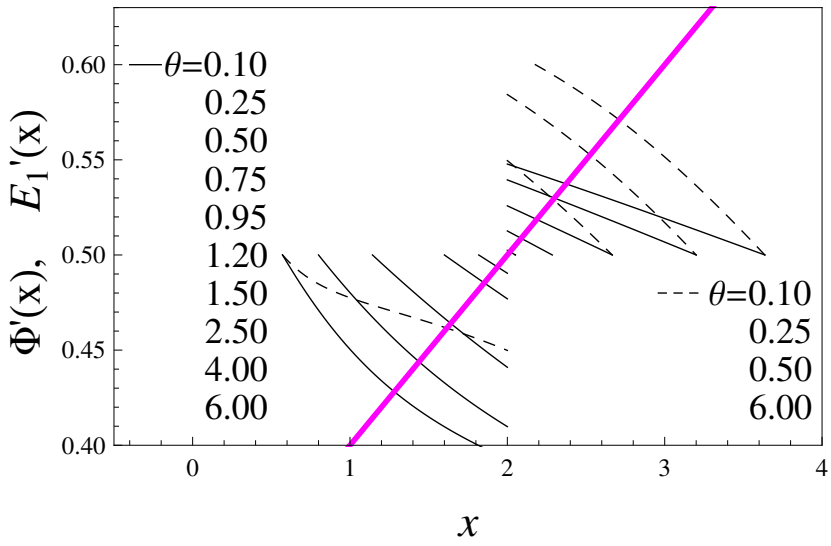

Figure 3.3: Full adaptation scenario ($\gamma < 0$). $\Phi'(x)$ with $f_{ij}[y] = e^{\beta_{ij}y/2}$ for varying β_{31} and fixed β_{32}; cf. Fig. 1.1. We assume $\hat{E}_3(x) = x/2$, $\hat{E}_2(x) = x - 2$, and (3.20) holds for $x \in (\frac{4}{1+\theta}, 2)$. Normal (resp. dashed) curves: $\Phi'(x)$ for x that support (3.20) under $\beta_{32} = 1$ (resp. $\beta_{32} = 3$) and various $\theta = \beta_{31}/\beta_{32}$, as indicated from the top to the bottom in the left (resp. right). The magenta (bold) curve shows $-E_1'(x)$, where for the considered range of x, $E_1'(x) = -0.1(x-2) - 0.5$. Adaptation conditions (3.23) hold for all curves, and for all β_{31}, β_{32}.

now $\Phi'(x)$ is very robust with respect to changing β_{31} and β_{32}, i.e. the adaptation is achieved for *all* β_{31} and β_{32} (excluding a $|\gamma D|$-dependent vicinity of $\beta_{31} = \beta_{32}$). Though the choice of $E_1(x)$ is more flexible than for the previous restricted adaptation scenario, it cannot belong to the set (3.13) of harmonic functions. For $\gamma D < 0$, x should change in a bounded domain; otherwise for the natural shape of energies ($E_i(x) \to \infty$ for $x \to \pm\infty$) one gets a non-normalizable $p(x)$ in (3.16).

Chapter 4

Statistical tests for MIXMAX pseudorandom number generator [1]

4.1 Introduction

In recent years, there is a growing interest on PRNGs in different branches of physics and not only. A good PRNG is important to have guaranteed results of Monte Carlo(MC) methods. There are many software packages for MC simulations where PRNGs are the central components. Among these packages one can mention the Geant4/CLHEP [140], a widely used simulation toolkit in HEP for modeling the passage of elementary particles through matter, also used for medical and space science simulations.

PRNGs are also crucial in Markov Chain Monte Carlo (MCMC) methods which are used for sampling from desired probability distribution by constructing Markov chain on state space whose stationary distribution is of interest [1, 131, 132]. Uniform PRNGs play a central role in constructing such Markov Chains. Most of MCMC algorithms are developed within random walk models. A widely used example of random walk Monte Carlo method is Metropolis–Hastings algorithm [131–134] which is also included in the list of the top 10 algorithms [241]. MCMC methods are mainly used for sampling from large dimensional spaces

[1] The results considered in this chapter are published in Ref. [141].

and computing multidimensional integrals. For example, in statistical mechanics, one needs to compute thermal averages of quantities, such as the total energy, magnetization, etc. by performing multidimensional integration or summation over configuration space. However, the total number of configurations can be very large, e.g. in 3-dimensional Ising model the number of spin configurations with particles at n^3 lattice sites is 2^{n^3}. In thermodynamic equilibrium the probabilities of occurring each configuration is represented by Boltzmann distribution. Thereby having samples drawn from Boltzmann distribution one can compute expectation values of thermodynamic quantities.

The necessity to have large amounts of simulated data imposes a strict requirements on PRNGs, such as statistical properties of generated numbers, swiftness in number generation, replicability, lengthiness of generated random cycle and independence of produced random numbers. To address these challenges the renewed version of MIXMAX PRNG [136, 137] based on Anosov C-systems and Kolmogorov K-systems has been introduced in [2, 138, 139]. The MIXMAX is matrix-recursive PRNG and it has been shown that the properties of the MIXMAX generator is improved with increasing the size N of MIXMAX matrix [2]. The period of MIXMAX is also increased with increasing N and it can be reach up to 10^{57824}, note that the period of commonly used version of Mersenne Twister based on Mersenne prime has the period of $2^{19937} - 1$.

While having a long period, however statistical properties and time characteristics of PRNGs are crucial to consider a generator "good" or "bad". In this paper we will present the results of the statistical tests performed with the matrix size of $N = 256$ which is considered to be a default dimension of MIXMAX matrix with flexibility to be further increased.

4.2 Visual demonstration

We can reveal the defect of uniform PRNGs simply plotting random points in high-dimensional Euclidean space, if these points form lattice structure then to a first approximation we can say that PRNG has defects in generating random points since the space is not filled uniformly. The Fig.1 shows the comparison of MIXMAX with the Linear Congruential

Generator(LCG), which is known to be defective PRNG. In contrast to LCG MIXMAX does not form lattice structure. We obtain these figures by generating two $U(0,1)$ random number sequences and assigning a point in two-dimensional space.

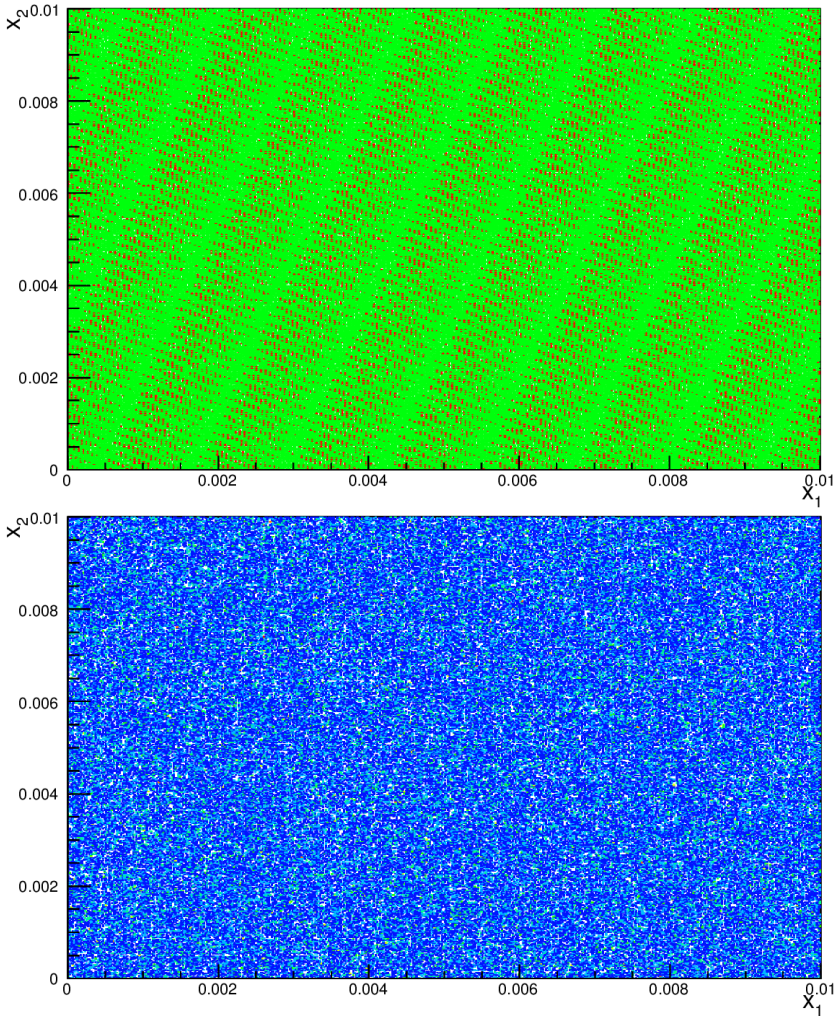

Figure 4.1: Random points in two-dimensional space generated by LCG(up) and MIXMAX(down)

4.3 Statistical testing with TestU01

Most of PRNG algorithms produce numbers uniformly distributed in the interval of $(0,1)$, hence PRNGs should pass statistical tests of uniformity. Many empirical statistical testing packages implement tests for these purposes, some of which are [242–245]. Most of the statistical tests implemented in these packages are discussed in Knuth's book [246], e.g. the

package [243] implements mainly tests of Knuth. Currently, one of the well known tools for statistical testing is TestU01 software library which provides implementations of the empirical statistical tests for uniform PRNGs. It contains more than 160 different empirical tests and offers several batteries of tests including the most powerful one, i.e. the 'Big Crush'. When a specific statistical test is applied to random numbers produced by PRNG the p-value of the test is printed as a measure of deviation from null-hypothesis, which in our case is uniform distribution of random numbers. In comparison with other libraries TestU01 is more flexible and efficient, and it can deal with larger sample sizes and has wider range of test statistics than other libraries.

In the Table 4.1, the outcome of TestU01 BigCrush suite applied on MIXMAX, Mersenne Twister and LCG is stored by using 64-bit computer with $Intel\ Core\ i3\ -4150$ processor of clock speed $3.50 \times 4\,GHz$.

Table 4.1: TestU01 BigCrush suite results

PRNG	Total CPU time	BigCrush	Failed test'(s) p value
MIXMAX	$2h\ 43m\ 51s$	*All tests were passed*	—
Mersenne Twister	$3h\ 19m\ 27s$	3	$0.9990, 1 - 10^{-15}$
LCG	$3h\ 30m\ 33s$	22	$< 10^{-300}$

As we can see from the table the MIXMAX passes the same test suite faster than Mersenne Twister and does not fail any test. TestU01 test suite has been applied to Ranlux PRNG with its modifications Ranlux24, Ranlux48. It is observed that Ranlux though having good statistical properties is very slow at generating random numbers. Comparing with MIXMAX Ranlux24 is 10 times slower and Ranlux48 is 17 times slower. This fact makes it not convenient for the use in generation of large amount of random numbers.

4.4 Kolmogorov-Smirnov tests

Kolmogorov-Smirnov (K-S) test is one of the powerful tools that can be used to examine the statistical features of PRNGs.

Though one-dimensional (1D) K-S test is already implemented in TestU01, we perform K-S test independently for various parameters of sample size (n) and extract the distribution of K-S test statistic. The idea behind the test is to calculate maximum distance between expected Cumulative Distribution Function(CDF) $F(x)$, $F(x) = Pr(X \leq x)$ and measured or Empirical Cumulative Distribution Function(ECDF) $F_n(x)$ of n data points

$$F_n(x) = \frac{1}{n}(number\ of\ x_i \leq x) \tag{4.1}$$

The null hypothesis H_0 is whether the sample of n random numbers comes from expected distribution $F(x)$ or not. If data comes from $F(x)$, then the strong law of large numbers provides $F_n(x) \to F(x)$, as $n \to \infty$. The latter is strengthened by the Glivenko-Canteli (G-C) theorem [247], which states that under H_0 hypothesis

$$Pr(\lim_{n\to\infty} sup_x\, |F_n(x) - F(x)| = 0) = 1 \tag{4.2}$$

Hence the difference between CDFs can be used as a measure of agreement between a data and a given distribution. There are several statistical tests based on (G-C) theorem known as Cramer-von Mises tests [248, 249]. The key feature of tests based on G-C theorem is that their distributions are independent of the hypothesized models under H_0 when data sample is large.

The one dimenisonal(1D) K-S test is defined as follows:

$$D_n = sup_x\, |F_n(x) - F(x)| \tag{4.3}$$

Under null hypothesis the distribution of $\sqrt{n} \cdot D_n$ converges to Kolmogorov distribution for sufficiently large n when $F(x)$ is continuous [250]

$$\lim_{n\to\infty} Pr(\sqrt{n}\cdot D_n \leq x) \equiv K(x) = 1 - 2\sum_{i=1}^{\infty}(-1)^{i-1}e^{-2i^2x^2} \tag{4.4}$$

It is of interest to note that for small values of n Kolmogorov distribution is not adequate, but there is a way to compute $p-value$ for randomly produced D_n [250].

In Fig. 4.2 the normalized histogram of $\sqrt{n}\cdot D_n$ data points for MIXMAX and Mersenne Twister is presented and compared with Probability Density Function (PDF) of Kolmogorov distribution:

$$f(x) = K'(x) = 8x\sum_{i=1}^{\infty}(-1)^{i-1}i^2e^{-2i^2x^2} \tag{4.5}$$

The histograms in Fig.(4.2–4.7) are normalized to unity dividing each bin entry by the product of sample size and bin width $(n\cdot width)$. Visual comparison shows that under H_0 the distribution of $\sqrt{n}\cdot D_n$ follows the PDF of Kolmogorov distribution. Due to fast convergence of the series of partial sums in Eqs.(4.4,4.5) it suffices to take limited number of terms in the sum, e.g. the first 100 terms are enough.

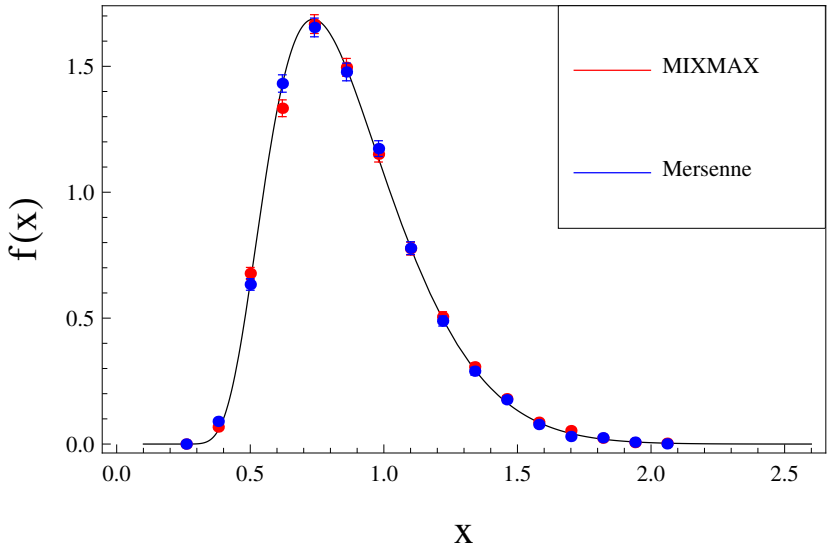

Figure 4.2: Distribution of $\sqrt{n}\cdot D_n$ for MIXMAX and Mersenne Twister. The size of a samples is $n = 10^8$ and the number of different replicas is 10^4. The black curve is the PDF of Kolmogorov distribution.

Two-level tests can be used on K-S test to give evidence of visual coincidence on Fig. 4.2. For this purpose chi-square test(next section) is applied. It has been checked that both distributions agree with theoretical expectation (black curve) in 95% Confidence Level (CL).

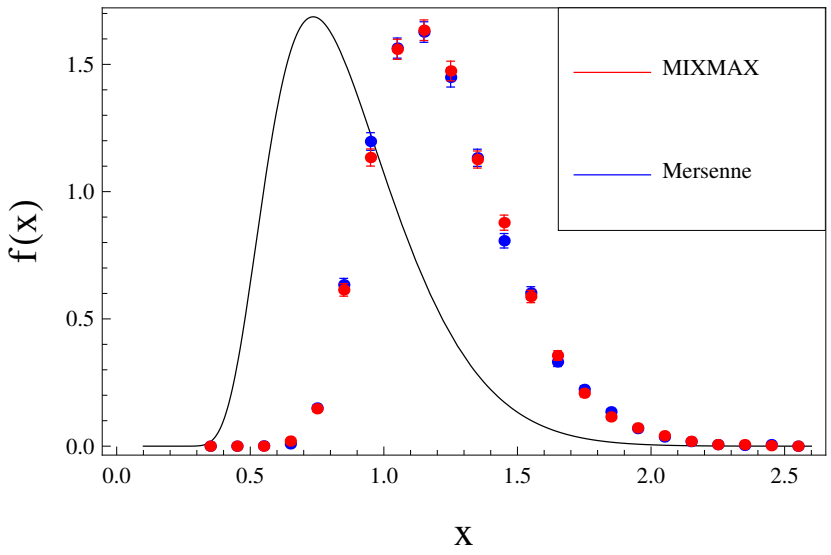

Figure 4.3: Distribution of $\sqrt{n} \cdot D_n$ for MIXMAX and Mersenne Twister for 2-dimensional case. The size of a samples is $n = 10^3$, note that n is the total number of random points in 2-dimensional space, i.e. 10^6 numbers are generated by PRNGs, the number of different replicas is 10^4. The black curve is the PDF of Kolmogorov distribution. The shape of histogram shows clear shift from Kolmogorov distribution. MIXMAX and Mersenne Twister that have been used in these studies give almost the same distributions for $\sqrt{n} \cdot D_n$ which seems to be independent from dimension of Kolmogorov-Smirnov test.

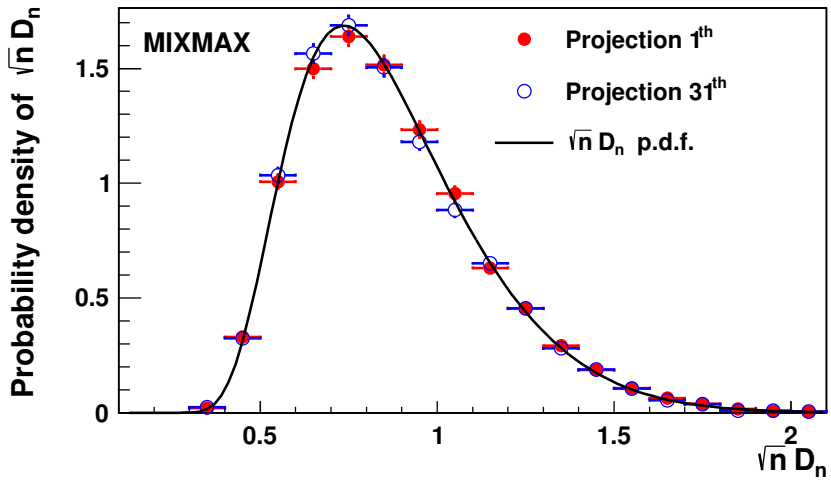

Figure 4.4: Distribution of $\sqrt{n} \cdot D_n$ for the 1^{th} and 31^{th} projections of MIXMAX RNG in comparison with the theoretical expectation (black curve). The sample size is $n = 10^8$ and for each projection 10^4 different replicas are generated.

In multidimensional K-S test one have d-dimensional data ($d \geq 2$) and to test H_0 it is needed to compare d-variate ECDF with the hypothetical d-variate CDF. The complication in multidimensional case is caused by the ambiguity in definition of the CDF since there are

$2^d - 1$ independent ways of defining CDFs.

There have been proposed different ways to calculate the multidimensional K-S statistic [251], [252]. In [251] four quadrants around all combinations (x_i, x_j) of data points is considered and D is taken as the maximum of 4 differences between CDFs over all quadrants. Therefore this idea makes the test statistic independent of ordering the data. The number of all pairs (x_i, x_j) for N points is equal to N^2, therefore to calculate D_n one needs to compute differences between CDFs in $3N^2$ quadrants(the probability for fourth quadrant is found from normalization). This method suffers from the computing time when N is large. In [252], it is proposed to consider only observed points rather than all combinations thereby reducing the computational time by computing the differences for $3N$ quadrants only. It is possible to compute D_n with computationally higher efficiency introducing a binning technique applied to a continuous multidimensional data, i.e. discretizing the data space. The idea of binning technique is discussed in the next chapter whose results are published in [253]. In [254] the algorithm for 2D K-S test is presented when only one CDF from all possible configurations is taken into account. As a result of it, the procedure used to compute D evaluating the difference of CDFs is reduced to a small number of data points. In our studies we have extended the standard definition of one dimensional cumulative distribution to its two dimensional "analogue" and computed K-S statistic using algorithm presented in [254](see Fig. 4.3). Three-dimensional extension of K-S test is presented in [255], where it is considered 8 CDFs, and using MC techniques the table of critical values are also presented in this paper.

To detect possible non-uniformities in the multidimensional random sequences of MIX-MAX PRNG, an arbitrary selected projection has been checked via Kolmogorov-Smirnov test and the results are compared with that of the first projection. In Fig. 4.4, the probability density distributions of Kolmogorov-Smirnov statistic are presented for the 1^{th} and the 31^{th} projections and the comparison is provided with the expected distribution.

4.5 χ^2 tests

The chi-square χ^2 test is one of the famous statistical tests which is met in many applications when one deals with grouped or binned data. The Chi-Square test is applied to categorical sample distributions unlike the K-S test which using each random point compares continuous sample distributions with hypothesized ones. The χ^2 statistic has the following form [246, 256]

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}, \tag{4.6}$$

where O_i is the observed number of data points in ith bin and $E_i = np_i$ is the expected number of data points falling into ith bin, here p_i is the probability that observation falls into ith bin. To apply the test to PRNGs [0,1] interval is divided into k bins and the χ^2 statistic is computed noting that $p_i = 1/k$. If H_0 is true then statistics defined in (4.6) computed for random samples follows chi-squared distribution with $\nu = (k-1)$ degrees of freedom

$$g_\nu(y) = \frac{2^{-\frac{\nu}{2}} e^{-\frac{y}{2}} y^{\frac{\nu}{2}-1}}{\Gamma(\frac{\nu}{2})}, \tag{4.7}$$

where $\Gamma(\nu)$ is gamma function. It is useful to introduce new random variable $x = \frac{y}{\nu}$ and consider the PDF of x denoted as $f_\nu(x)$. This enables to get rid of small numbers in PDFs when ν is big. Since

$$1 = \int f_\nu(x)dx = \int g_\nu(y)dy \tag{4.8}$$

it follows that

$$f_\nu(x) = \nu g_\nu(\nu x) = \frac{2^{-\frac{\nu}{2}} \nu^{\frac{\nu}{2}} e^{-\frac{x}{2\nu}} x^{\frac{\nu}{2}-1}}{\Gamma(\frac{\nu}{2})} \tag{4.9}$$

The new random variable introduced in (4.9) is called reduced chi-square. The distribution of reduced chi-square for MIXMAX and Mersenne Twister is shown in Fig. 4.5 in comparison with theoretical expectation.

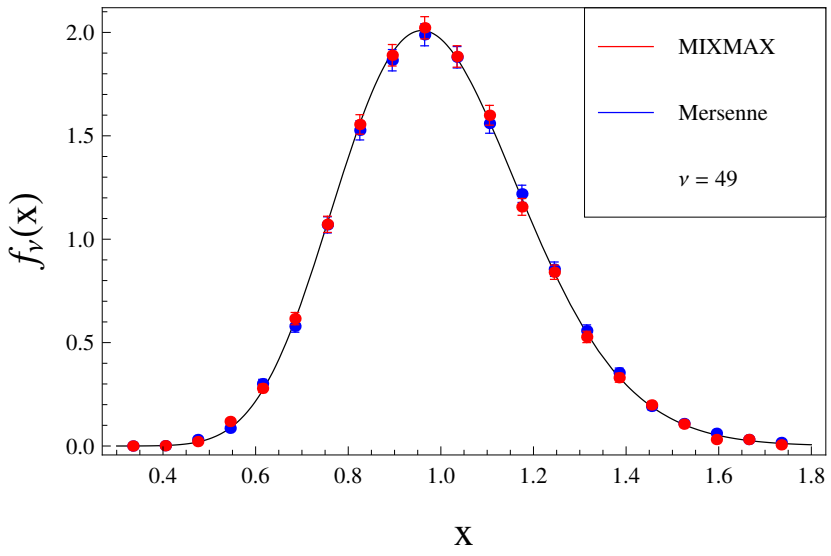

Figure 4.5: Comparing denisty histogram of reduced chi-square for MIXMAX and Mersenne Twister with the PDF in (4.9). The size of a samples is $n = 10^6$ and the number of different replicas is 10^4.

4.6 Serial tests

The serial test also known as chi-square test of independence is multidimensional analogue of chi-square test which checks the independence between two or more random variables [246, 257]. When serial test is applied to PRNGs one divides random sequence into groups of non-overlapping d-tuples $(x_{id}, x_{id+}, ..., x_{id+k-1})$, where $i = 1, 2, ..., \frac{n}{d}$, hence the elements of d-tuple are considered as realizations of d random variables and the relationship between them is of interest . If x_is are $U[0, 1]$ random variables then k-tuples are uniformly distributed in $[0, 1]^d$. To check this each dimension of unit hypercube is divided into k bins and the data of d-tuples is binned into $[0, 1]^d$. Now chi-square statistic (4.6) is applied to this data comparing the number of observations falling in each sub-hypercube with theoretical expectation: $E_{i,j,...,d} = np_{i,j,...,d}$, where the joint probability $p_{i,j,...,d}$ of d-dimensional data point to fall into $(i, j, ..., d)$ sub-hypercube is the product of probabilities of each individual coordinate to fall into appropriate bin, which is the condition of independence.

$$p_{i,j,...,d} = \prod_{n=1}^{d} p_n = \left(\frac{1}{k}\right)^d \tag{4.10}$$

Unlike the non-overlapping tuples, the overlapping d-tuples of random sequence fall on neighboring parallel planes. The largest distance between adjacent parallel hyperplanes is called the spectral test statistic [246, 258–262]. If the largest distance is small then it implies that overlapping d-tuples are more uniformly distributed in unit hypercube, therefore PRNG is considered good.

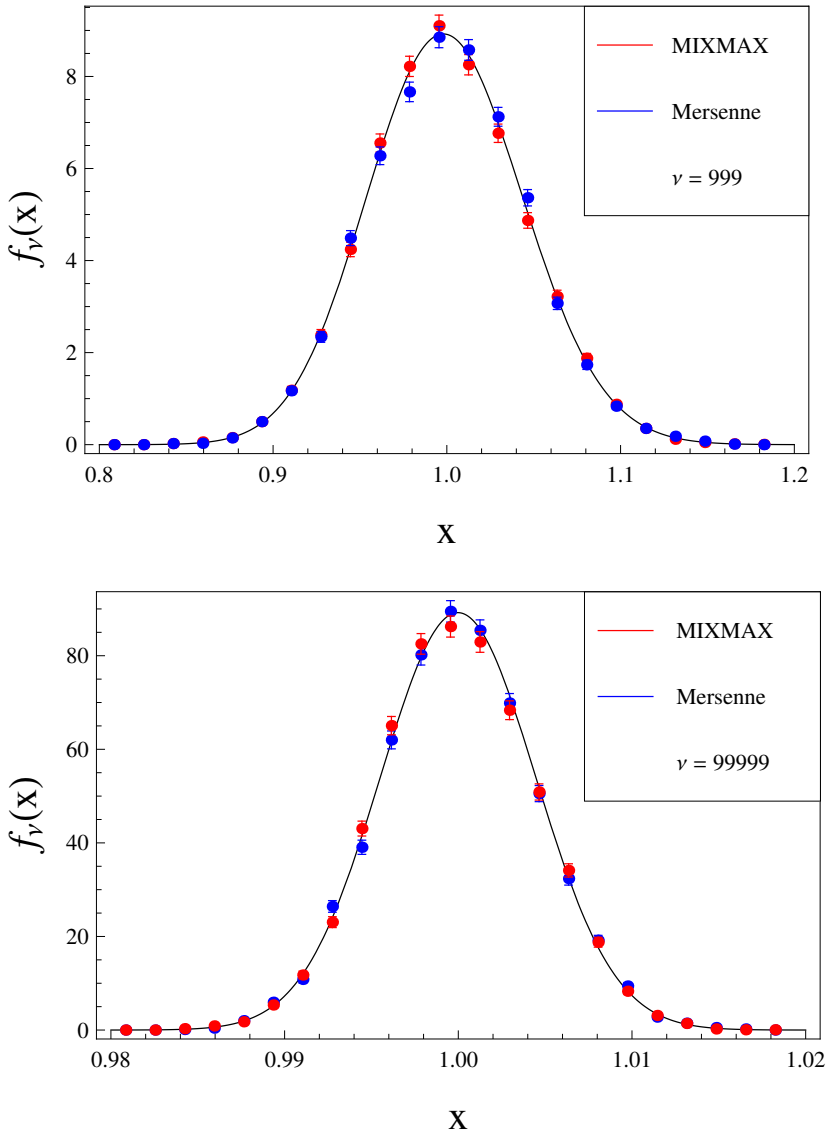

Figure 4.6: Comparing denisty histogram of reduced chi-square with the test distribution (4.9) for 3-dimensional(up) and 5-dimensional(down) cases.

The Fig. 4.6 shows $d = 3$ and $d = 5$ dimensional cases of serial test, where each dimension is divided into $k = 10$ bins and the histogram of reduced chi-square test statistics is compared with the distribution of (4.9) with appropriate degrees of freedom. The reduced chi-square distribution reveals no significant distinction between MIXMAX and Mersenne Twister. Note

that the serial test here is applied to single random stream and measures correlations between adjacent random tuples. This test can be also applied to different streams to check the independence between them.

4.7 Parallel streams of MIXMAX

All tests described in previous sections use one stream generated by PRNG. However, in multiprocessor stochastic computations it is important to have uniformly distributed and statistically independent simultaneous random streams partitioned across the processors [263–267]. Different parallelisation approaches of PRNGs have been studied in literature [267,269–271]. One trivial technique for parallelisation is to take random seeds on each processor, but since every PRNG has finite number of states, one should be careful in order to avoid possible overlapping between different streams. MIXMAX has very large state space, therefore even taking random seeds on each processor does not affect the independence between multiple streams. Another approach is to take single sequence and partition it into different processors. MIXMAX provides skipping-ahead algorithm which enables to skip forward by large amount of numbers in sequence, this technique guarantees the non-collision of partitioned streams [269]. The check for randomness of each individual stream can be done via the standard chi-square or K-S tests. The test of independence of multiple streams can be done via parallel version of serial test simply forming d-tuples of random numbers taken from each of d streams at a time. However, it is not practical to test empirically all random streams when the period is very large, different techniques for testing parallel streams can be found in [267].

The serial test up to dimensions $d = 7$ has been performed and it is observed that multiple streams of MIXMAX are statistically independent which is also guaranteed by underlying theory of MIXMAX.

One can analyze the independence and uniformity of parallel streams using the fact that the sum of n independent $U[0, 1]$ random variables follow Irwin–Hall distribution of order n [272].

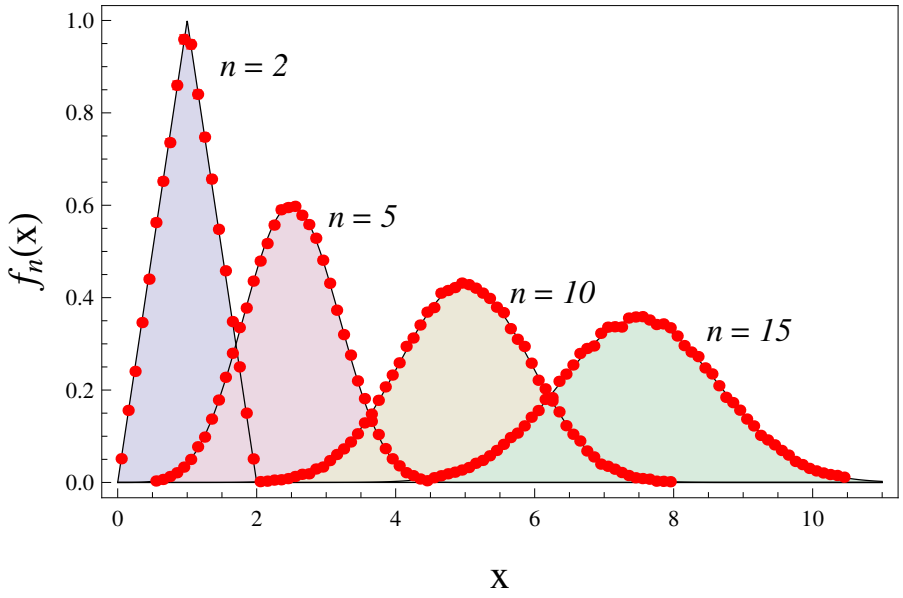

Figure 4.7: The distribution of Irwin-Hall of order (2,5,10,15) compared with density histogram of data.

Under H_0 if each stream of PRNG is generated from uniform distribution then the random sequence resulted from element-wise addition of multiple streams has Irwin–Hall distribution. The Irwin–Hall distribution has the following form

$$f_n(x) = \frac{1}{2(n-1)!}\sum_{i=1}^{n}(-1)^i\binom{n}{i}(x-i)^{n-1}sgn(x-i), \tag{4.11}$$

where $sgn(x-i)$ is $sign$ function. When $n = 2$, then (4.11) reduces to the well known triangular distribution

$$f_2(x) = \begin{cases} x, & 0 \le x < 1 \\ 2-x, & 1 \le x \le 2 \end{cases} \tag{4.12}$$

Table 4.2: p-values from chi-square test

Test	n = 2	n = 5	n = 10	n = 15
***p*-value**	0.76	0.97	0.39	0.37

In Fig. 4.7 visual comparison of the data with Irwin–Hall distribution is shown, where up

to 15 random streams is taken to form the sum. To check visual consistency of the histogram and model prediction in Fig.(4.7) chi square test is applied for comparison with Irwin-Hall distribution. The Table 4.2 represents p-values of chi square statistic.

Chapter 5

The Optimal Approach for Kolmogorov-Smirnov Test Calculation in High Dimensional Space [1]

5.1 Introduction

One of the most important tasks nowadays is the numerical experiments on super computers with the modeling of the sophisticated physics system. The Monte Carlo method is used for solving the problems where the high dimensional integration is involved. To use the Monte Carlo method for analysis of the high energy physics experimental and theoretical problems we have to solve the problem of the quality of the pseudo-random generators, which should have a strong statistical feature, also a large period of sequences and a high speed of generating of the pseudo-random number. The K-S test is used to compare how well the ECDF of a sample fits the CDF of the reference distribution by computing the maximum distance between these two functions. The K-S test is applied to exactly continuous data, and for dimensions $d \geq 2$ we have to compute the distance on infinite points then take a maximum, because when $d \geq 2$ for CDF and ECDF we have d-dimensional surfaces.

[1] The results considered in this chapter are published in Ref. [253].

Several algorithms for computing the two-dimensional K-S test were proposed [251, 252, 254, 255]. We propose another approach for 2 and higher dimensions. Our method is based on discretization of the space introducing a binning technique applied to continuous multidimensional data. This technique does not correspond to the usual K-S test principle, but we show that it is possible to compute K-S statistics very precisely with the quite computationally efficient algorithm taking minimal bin number. Here our K-S test results are obtained using the well-tested uniform Mersenne Twister PRNG [3], which is included in CERN library [268]. In the proposed paper we describe a new technique, which allows to reduce essentially the number of the bins and correspondingly decrease the needed time and memory used.

5.2 Formalism

In one-dimensional (1D) case to check whether the sample of n random numbers come from uniform distribution in the $[0, 1]$ range or not we need to compute K-S test with the CDF of uniform distribution

$$F(x) = x \tag{5.1}$$

Usual (unbinned) approach in 1D case looks as follows:

(i) n numbers should be sorted in increasing order $\{x_1, x_2, \ldots, x_n\}, x_i \leq x_{i+1}$

(ii) D_n is computed in the following way:

$$D_n = \max_{0 \leq x \leq 1} |F_n(x) - x| = \max_{1 \leq i \leq n} \left\{ \left| \frac{i}{n} - x_i \right|, \left| \frac{i-1}{n} - x_i \right| \right\}, \tag{5.2}$$

where $F_n(x)$ is ECDF of n data points

$$F_n(x) = \frac{1}{n}(number\ of\ x_i \leq x) = \frac{k}{n}, \quad x = x_1, \ldots, x_n \tag{5.3}$$

k in (5.3) is the number of random points out of n with $x_1, \ldots, x_k \leq x$.

The unbinned approach can be also applied for 2D case following the algorithm described in [254], which has been used for the performed studies.

In the 1D binned approach the $[0,1]$ interval is divided into b bins, then n random numbers are distributed in b bins according to their values. The ECDF is again defined as:

$$F_n(x) = \frac{k}{n} \tag{5.4}$$

The difference is that here x takes values of bin edges: $x = \frac{i}{b},\ i = 1, \ldots, b$. For example, if $b = 10$, we have the bins edges like $\{0.1, 0.2, \ldots, 0.9, 1\}$ and k is the number of points whose values are less than equal to the value of bin edge. Therefore, the advantage of the method is that if n is quite large then ECDF can be defined with a relatively small number of points $(b < n)$. Now D_n is computed as follows

$$\begin{aligned} D_n = \max_{0\le x\le 1} |F_n(x) - x| &= \max_{1\le i\le b} |F_n(x_i) - x_i| \\ &= \max_{1\le i\le b} \left| \frac{\sum_{j=1}^{i} Y_j}{n} - \frac{i}{b} \right|, \end{aligned} \tag{5.5}$$

where Y_j is the number of observations that do fall into jth bin.

The 1D binned method can be generalized for higher dimensions. In two-dimensional case (when only one CDF from four possible configurations is taken into account, see the section of Kolmogorov-Smirnov tests in the previous chapter) the corresponding expression for D_n is

$$\begin{aligned} D_n = \max_{\substack{0\le x_1\le 1 \\ 0\le x_2\le 1}} |F_n(x_1, x_2) - x_1 x_2| &= \max_{\substack{1\le i\le b \\ 1\le j\le b}} |F_n(x_i, x_j) - x_i x_j| \\ &= \max_{\substack{1\le i\le b \\ 1\le j\le b}} \left| \frac{\sum_{m=1}^{i} \sum_{n=1}^{j} Y_{mn}}{n} - \frac{ij}{b^2} \right|, \end{aligned} \tag{5.6}$$

where Y_{mn} is the number of observations that fall into category $\{m, n\}$. Here it was used the usual definition of joint CDF(ECDF) of two independent uniform random variables, i.e. $F_{XY}(x, y) = Pr(X \le x, Y \le y) = xy$. Note that each axis is divided into b bins.

The generalization of the formula for any dimension d has the form

$$D_n = \max_{\substack{0\le x_1\le 1\\ \dots\\ 0\le x_d\le 1}} |F_n(x_1,\dots,x_d) - x_1\dots x_d| = \max_{\substack{1\le i_1\le b\\ \dots\\ 1\le i_d\le b}} |F_n(x_{i_1},\dots,x_{i_d}) - x_{i_1}\dots x_{i_d}|$$
$$= \max_{\substack{1\le i_1\le b\\ \dots\\ 1\le i_d\le b}} \left| \frac{\sum_{m_1=1}^{i_1}\dots\sum_{m_d=1}^{i_d} Y_{m_1\dots m_d}}{n} - \frac{i_1\dots i_d}{b^d} \right| \tag{5.7}$$

5.3 Binning technique and Kolmogorov distribution

We aim to show that the PDF (or mean) of K_n computed via binning technique with small number of bins ($b < n$) is close to the exact PDF (or mean of exact PDF) given by Kolmogorov distribution in 1D case [250]

$$f(x) = 8x\sum_{i=1}^{\infty}(-1)^{i-1}i^2e^{-2i^2x^2},$$
$$< K_n >= \int_0^{\infty} xf(x)dx = 0.8687\dots \tag{5.8}$$

Now visual comparison of Fig.(5.1) confirms that binning technique approach is reliable approximations since distribution of K_{10^5} computed $b = 10^4$ is close to the Kolmogorov distribution.

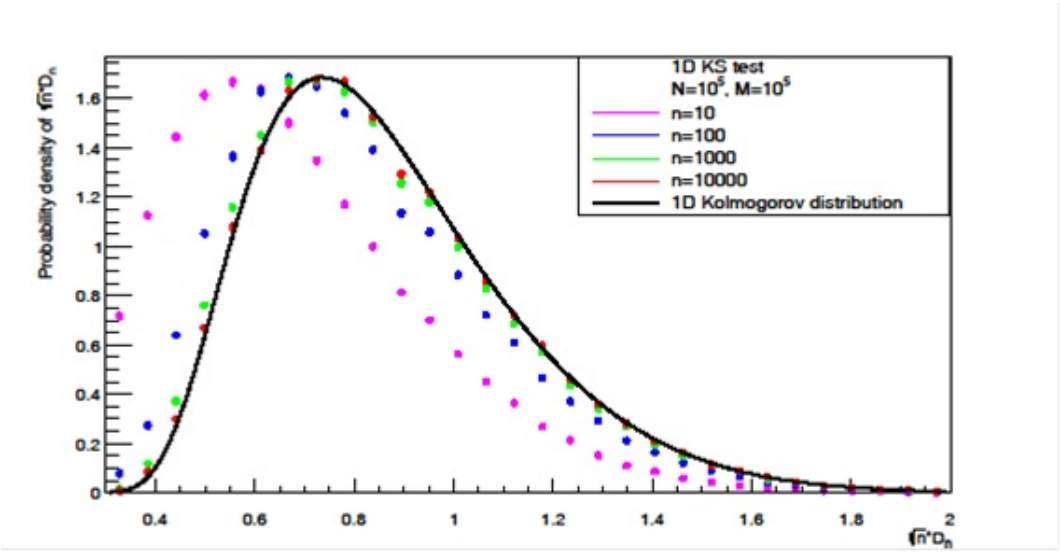

Figure 5.1: Binning technique approach vs Kolmogorov distribution. Following to notification on this figure the size of a samples is $N = 10^5$, the number of bins n (in the text N is denoted as n and n is denoted as b), and the number of different replicas is $M = 10^4$. The black curve is the PDF of Kolmogorov distribution.

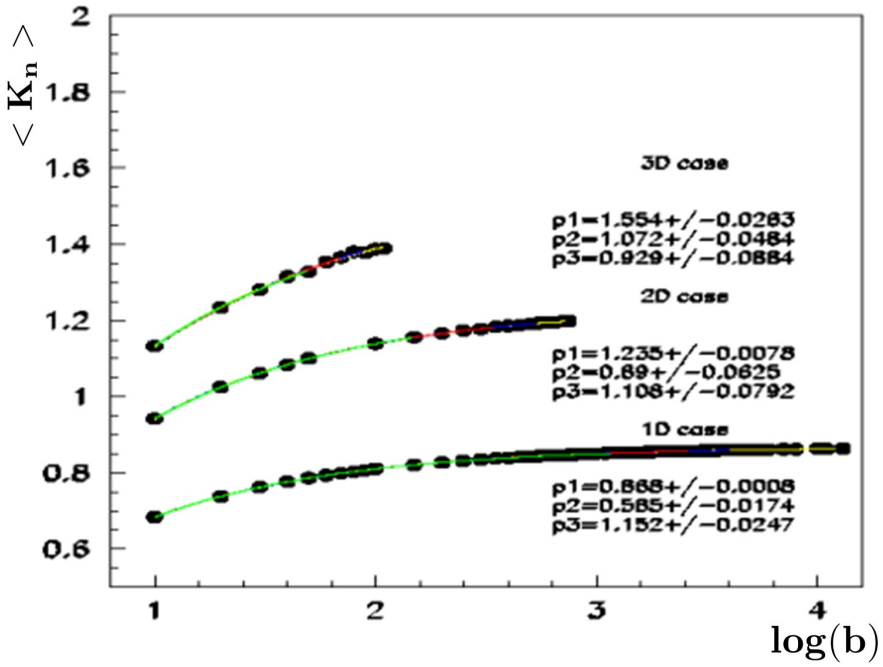

Figure 5.2: Binning technique approach vs Kolmogorov distribution. The size of a samples is $n = 10^5$, the number of bins is 10^4, and the number of different replicas is $M = 10^4$. The black curve is the PDF of Kolmogorov distribution.

In Fig.(5.2) the dependence of K_n on the number of bins is presented for 1, 2 and 3D cases. Here one can recognize the clear dependence on the number of bins (b). Taking into account the obtained saturating shape for such dependencies we performed the fit with the following fitting function:

$$f_{fit}^{\boldsymbol{p}} = p_1 + p_2 e^{-p_3 x} \tag{5.9}$$

The meaning of the parameter p_1 corresponds to the saturation level, which can be achieved in particular with quite large value of b, i.e. for fixed n $b \Rightarrow n$. Although, using very high values of the bin number is not practical in the sense of the needed CPU time and memory, especially in case of higher ($d > 4$) dimensional space. We can compute $< K_n >$ using a limited number of bin points, then estimate the saturation parameter b, then make

a correction for the average value of $< K_n >$ estimated with e.g. $b = 10$, introducing the correction factor: $C_{10} = p_1 - < K_n(b = 10) >$, this approach seems to be very interesting and effective in order to realize the behaviour of $< K_n >$ estimation in case of high dimensional space. Also the exponential coefficient p_3 which regulates the speed of convergence to the saturation level, can be used to check the quality of different PRNGs. It should be noted, that the used fitting function is quite stable in respect to the variation of the fitting point numbers. In 3D case with the total number of points ($n_{total} = 11$), this variation was 11, 9, 7, 5. In 2D case with the $n_{total} = 19$: 19, 15, 11, 7, and in 1D case with $n_{total} = 75$: 75, 55, 35, 15. This is an important feature of the functional form used for a fit, because in case of essentially higher dimensions one can compute with the binned method a limited number of points, perform the fit and estimate the saturation level.

Conclusion

It was shown that there is a mechanism by which the no-pumping theorem|time-averaged probability currents nullify in the Floquet regime|holds for the destination transition rates (1.6) (which hold the detailed balance condition). A sufficient condition for this validity is that the external time-periodic fields acting on the stochastic system hold the time-symmetry (1.40). Similar time-symmetry conditions (together with the space symmetry of the external potential) govern the current-generation regimes of various ratchet models; see [164] for a recent review. It should be interesting to understand in more detail possible relations betweem the no-pumping theorem and the (no) current-generation in ratchets; this is left for future work. In the regime, where the no-pumping theorem holds exactly for the destination rates, there is an approximate no-pumping theorem that holds for the physically pertinent Fokker-Planck rates.

It was studied a class of non-equilibrium stationary states generated by two thermal baths at different temperatures. Three conditions were assumed:

(i) Variables of the system have different characteristic times.

(ii) External fields act only on the slow variable.

(iii) The transition rates of the slow variable are such that (under time-scale separation) it holds an effective detailed balance. Examples of the effective detailed balance are the activation transition rate for an arbitrary topology of connections, and arbitrary rates for a tree-like topology.

Conditions *(iii)* are similar to those governing the no-pumping theorem.

Under conditions *(i)–(iii)* there exists a non-equilibrium free energy. It describes quasi-

stationary isothermal processes, and is defined via the potential of the work done via a slow variation of external parameters. As compared to its equilibrium counterpart, the non-equilibrium free-energy can be an increasing function of the bath temperature. As a physical demonstration of this feature we studied the free-enery cost of the isothermal cooling (i.e. entropy decrease via slow external fields): this cost can decrease if the cooling starts from a higher temperature. It is interesting to compare this finding with the called Mpemba phenomenon: given two samples of water that are identical except of their initial temperature, the initially hottter samples cools (and freezes) quicker than the colder one. This can have relatively straightforward physical explanations, e.g. some part of the hotter water can evaporate thus decreasing its amount and making its cooling easier. Mpemba phenomenon disappears once such scenarios are ruled out. To compare our finding with the Mpemba effect, we first of all note that our cooling is not spontaneous. We focus on an (isothermal) entropy reduction of the system by means of external fields. (Still our cooling process is slow and it leaves the system in its locally stationary state.) Hence we study not the speed of cooling, but its work-cost, as determined by the free-energy. The effect we found is impossible in equilibrium, because the equilibrium free energy is a decreasing function of the temperature. However, we show that this effect survives close to equilibrium. Hence the second major difference with the Mpemba phenomenon is that our effect is essentially out of equilibrium, as it relates to two different temperatures.

The existence of the free energy is facilitated in the (one-dimensional) continuous limit of the slow variable. Now time-scale separation comes out naturally, and there is a situation, where a non-equilibrium free energy exists for all driven parameters, i.e. assumption *(ii)* can be relaxed.

It was studied a model for an adaptive heat engine that can function under scarce or unknown resources. Several physical limitations for the adaptation concept were uncovered. They relate to the prior information available about the environment, i.e. whether both bath temperatures vary or one of them is fixed.

The detailed studies of newly released MIXMAX PRNG was performed. The empirical statistical tests of uniform random number generators implemented in TestU01 software li-

brary were applied to MIXMAX and other generators available in market. The main goal was to compare the features of MIXMAX with the most widely used Mersenne Twister. In addition to TestU01 (which gives p-value of statistical test when applied to random numbers produced by PRNG) some tests were performed independently for various parameters of sample size. The analysis reveal compatible statistical features of MIXMAX and Mersenne, but better time characteristics for MIXMAX in generating random numbers.

A new approach to optimize the Kolmogorov-Smirnov discrepancy calculation with the binned method for high dimensional spaces by means of reducing the used bin number is developed. This approach approach allows to extend the estimation of D_n to very high dimensional spaces with reasonable requirements to the CPU time and memory, and, thus, to provide very important tests of the PRNGs, which are used to apply the Monte Carlo method for solving of essentially high dimensional physics problems.

Bibliography

[1] A.T. Bharucha-Reid, *Elements of the Theory of Markov Processes and their Applications*,
Courier Corporation, (2012).

[2] K.G. Savvidy, *The MIXMAX random number generator*,
Comp.r Phys. Comm. **196**, pp. 161-165, (2015).

[3] M. Matsumoto and T. Nishimura, *Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator*,
TOMACS **8**, pp. 3-30, (1998).

[4] F. James, *RANLUX: A Fortran implementation of the high-quality pseudorandom number generator of Lüscher*,
Comp. Phys. Comm. **79**, pp. 111-114, (1994).

[5] R.C. Hilborn, *Sea gulls, butterflies, and grasshoppers: a brief history of the butterfly effect in nonlinear dynamics*,
Am. J. Phys. **72**, pp. 425-427, (2004).

[6] R. Srivastava, L. You, J. Summers and J. Yin, *Stochastic vs. deterministic modeling of intracellular viral kinetics*,
Journal of Theoretical Biology **218**, pp. 309-321 (2002).

[7] M. Scott *Applied stochastic processes in science and engineering*,
Citeseer, (2011).

[8] M. Pinsky and S. Karlin *An introduction to stochastic modeling*,
Academic press, (2010).

[9] N.G. van Kampen, *Stochastic processes in physics and chemistry*,
Elsevier, Amsterdam, (2007).

[10] L. J.S. Allen, *An Introduction to Stochastic Processes with Applications to Biology*,
CRC Press, Lubbock, Texas, (2010).

[11] P.C. Bressloff, *Stochastic processes in cell biology*,
Springer, (2014).

[12] H.C. Tuckwell, *Stochastic processes in the neurosciences*,
SIAM, (1989).

[13] G.J. Lord and C. Laing, *Stochastic Methods in Neuroscience*,
Oxford University Press, (2010) .

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

[15] F.S. Samaria and A.C. Harter, *Parameterisation of a stochastic model for human face identification*,
Proceedings of the Second IEEE Workshop on Applications of Computer Vision, Sarasota, pp. 138-142 (1994).

[16] J. Yamato, J. Ohya and K. Ishii, *Recognizing human action in time-sequential images using Hidden Markov Model*,
Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 379-385, (1992)

[17] M. Sonka, V. Hlavac and R. Boyle, *Image processing, Analysis and Machine Vision*,
Cengage Learning, (2014).

[18] J. Katz and Y. Lindell, *Introduction to modern cryptography*,
CRC press, (2014).

[19] R. Lande, S. Engen and B. Saether, *Stochastic Population Dynamics in Ecology and Conservation*,
Oxford University Press, (2003).

[20] G.P. Basharin, A.N. Langville and V.A. Naumov, *The life and work of A.A. Markov*,
Linear Algebra and its Applications **386**, pp. 3-26, (2004).

[21]] R. Newburgh, J. Peidle, W. Rueckner, *Einstein, Perrin, and the reality of atoms: 1905 revisited*,
American Journal of Physics **74**, pp. 478-481, (2006).

[22] H. Risken, *The Fokker-Planck Equation*,
Springer-Verlag, Berlin, (1984).

[23] N.G. van Kampen, *Remarks on non-Markov processes*,
Brazilian Journal of Physics **28**, pp. 90-96, (1998).

[24] L. Page, S. Brin, R. Motwani, and T. Winograd, *The Pagerank Citation Ranking: Bringing Order to the web*,
Technical report, Stanford University, (1998).

[25] K.A. Dill, S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*,
Garland Science, New York, USA, (2002).

[26] C. Tanford, *Physical Chemisry of Macromolecules*,
John Wiley & Sons, New York, USA, (1961).

[27] L. von Bertalanffy, *The theory of open systems in physics and biology*,
Science, **111**, pp. 23-29, (1950).

[28] G. Recordati, *A thermodynamic model of the sympathetic and parasympathetic nervous systems*,
Autonomic Neuroscience, **103**, pp. 1-12, (2003),

[29] D. Nickelsen, *Markov Processes linking Thermodynamics and Turbulence*,
arXiv preprint arXiv:1510.06281, (2015).

[30] H.M. Jaeger and A.J. Liu, *Far-from-equilibrium physics: An overview*,
arXiv preprint arXiv:1009.4874, (2010).

[31] R. Grimshaw, *Nonlinear Ordinary Differential Equations*,
Blackwell, Oxford, (1990).

[32] N.H. Martirosyan, *No-pumping theorem for non-Arrhenius rates*,
Chin. Jour. Phys. **55**, pp. 500-507, (2017).

[33] A. Leigh, J. K. Y. Wong, F. Dehez and F. Zerbetto, *Unidirectional rotation in a mechanically interlocked molecular rotor*,
Nature **424**, pp. 174-179, (2003).

[34] V. Balzani, A. Credi, F. M. Raymo and J. F. Stoddart, *Artificial molecular machines*,
Angew. Chem. Int. Ed. **39**, pp. 3348-3391, (2000).

[35] S. Rahav, *Validity of the no-pumping theorem in systems with finite-range interactions between particles*,
Phys. Rev. E **95**, pp. 012159, (2017).

[36] S. Rahav and J. Horowitz and C. Jarzynski, *Directed flow in non-adiabatic stochastic pumps.*
Phys. Rev. Lett. **101**, pp. 140602, (2008).

[37] R. D. Astumian, *Adiabatic operation of a molecular machine*,
Proc. Natl. Acad. Sci. **104**, pp. 19715-19718, (2007).

[38] V.Y. Chernyak and N.A. Sinitsyn, *Pumping-Restriction Theorem for Stochastic Networks*,
Phys. Rev. Lett. **101**, pp. 160601, (2008).

[39] C. Maes, K. Netočnỳ and S.R. Thomas, *General no-go condition for stochastic pumping*, J. Chem. Phys. **132**, pp. 06B612, (2010).

[40] D. Mandal and C. Jarzynski, *A proof by graphical construction of the no-pumping theorem of stochastic pumps*, J. Stat. Mech. **2011**, pp. P10006, (2011).

[41] D. Mandal, *Unification and new extensions of the no-pumping theorems of stochastic pumps*, EPL **108**, pp. 50001, (2014).

[42] D. Mandal and C. Jarzynski, *Hybrid Models of Molecular Machines and the No-Pumping Theorem*, J. Chem. Phys. **137**, pp. 234104, (2012).

[43] S. Asban and S. Rahav, *No-pumping theorem for many particle stochastic pumps*, Phys. Rev. Lett. **112**, pp. 050601, (2014).

[44] J. Ren, V. Y. Chernyak, and N. A. Sinitsyn, *Duality and fluctuation relations for statistics of currents on cyclic graphs*, J. Stat. Mech. **2011**, pp. P05011, (2011).

[45] R.K.P. Zia, L.B. Shaw, B. Schmittmann, and R.J. Astalos, *Contrasts between equilibrium and non-equilibrium steady states: computer aided discoveries in simple lattice gases*, Comp. Phys. Comm. **127**, pp. 23-31, (2000).

[46] N.P. Smith and E.J. Crampin, *Development of models of active ion transport for whole-cell modelling: cardiac sodium-potassium pump as a case study*, Prog. Biophys. Mol. Biol. **85**, pp. 387-405, (2004).

[47] G. Recordati and TG. Bellini, *A definition of internal constancy and homeostasis in the context of non-equilibrium thermodynamics*, Experimental physiology **89**, pp. 27-38, (2004).

[48] S. U. Pillai, T. Suel, and S. H. Cha, *The Perron-Frobenius theorem: some of its applications*, IEEE Signal Processing Magazine **22**, pp. 62–75, (2005).

[49] H. Qian and L. M. Bishop, *The Chemical Master Equation Approach to Nonequilibrium Steady-State of Open Biochemical Systems: Linear Single-Molecule Enzyme Kinetics and Nonlinear Biochemical Reaction Networks*, Int. J. Mol. Sci. **11**, pp. 3472-3500, (2010).

[50] I.R. Epstein and J.A. Pogman, *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*, Oxford University Press, Oxford, UK, (1998).

[51] P. Érdi and J. Tóth, *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models*,
Manchester Univ. Press, Manchester, UK, (1989).

[52] H. Qian, *Open-system nonequilibrium steady state: Statistical thermodynamics, fluctuations, and chemical oscillations*,
J. Phys. Chem. B **110**, pp. 15063–15074, (2006).

[53] H. Qian, *Phosphorylation energy hypothesis: Open chemical systems and their biological functions*,
Annu. Rev. Phys. Chem. **58**, pp. 113–142, (2007).

[54] R.K.P. Zia and B. Schmittmann, *Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states*,
J Stat Mech: Theor. Exp. **2007**, pp. P07012, (2007).

[55] R. Kubo, *The fluctuation-dissipation theorem*,
Rep. Prog. Phys. **29**, pp. 255, (1966).

[56] U. M. B. Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani, *Fluctuation-dissipation: response theory in statistical physics*
Phys. Rep.
textbf461, pp. 111-195, (2008).

[57] K.A. Dill and S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience*,
Garland Science, New York, (2010).

[58] U. Seifert, *Stochastic thermodynamics, fluctuation theorems, and molecular machines*,
Rep. Prog. Phys. **75**, pp. 126001, (2012).

[59] U. Seifert and T. Speck, *Fluctuation-Dissipation Theorem in Nonequilibrium Steady States*,
EPL **89**, pp. 10007, (2010).

[60] J. Casas-Vázquez and D. Jou, *Temperature in non-equilibrium states: a review of open problems and current proposals*,
Rep Prog Phys **66**, pp. 1937–2023, (2003).

[61] E. Ben-Isaac, Y. Park, G. Popescu, F. L. H. Brown, N. S. Gov, and Y. Shokef, *Effective temperature of red-blood-cell membrane fluctuations*,
Phys. Rev. Lett. **106**, pp. 238103, (2011).

[62] D. Loi, S. Mossa and L. F. Cuglia, *Effective temperature of active matter*,
Phys. Rev. E **77**, pp. 051111, (2008).

[63] W. F. Paxton, K. C. Kistler, C. C. Olmeda, A. Sen, S. K. St Angelo, Y. Y. Cao, T. E. Mallouk, P. E. Lammert, and V. H. Crespi, *Catalytic nanomotors: Autonomous movement of striped nanorods*,
J. Am. Chem. Soc. **126**, pp. 13424–13431, (2004).

[64] R. Raduenz, D. Rings, K. Kroy and F. Cichos, *Hot Brownian particles and photothermal correlation spectroscopy*,
J. Phys. Chem. A **113**, pp. 1674–1677, (2009).

[65] D. Rings, R. Schachoff, M. Selmke, F. Cichos and K. Kroy, *Hot Brownian motion*,
Phys. Rev. Lett. **105**, pp. 090604, (2010).

[66] D. Rings, M. Selmke, F. Cichos and K. Kroy, *Theory of hot Brownian motion*,
Soft Matter **7**, pp. 3441–3452, (2011).

[67] D. Chakraborty, M. V. Gnann, D. Rings, J. Glaser, F. Otto, F. Cichos and K. Kroy, *Generalised Einstein relation for hot Brownian motion*,
EPL **96**, pp. 60009, (2011).

[68] L. Joly, S. Merabia and J.-L. Barrat, *Effective temperatures of a heated Brownian particle*,
EPL **94**, pp. 50007, (2011).

[69] G. Falasco, M. V. Gnann, D. Rings and K. Kroy, *Effective temperatures of hot Brownian motion*,
Phys. Rev. E **90**, pp. 032131, (2014).

[70] R. Radünz, D. Rings, K. Kroy and F. Cichos, *Hot Brownian particles and photothermal correlation spectroscopy*,
J. Phys. Chem. A **113**, pp. 1674-1677, (2009).

[71] N. G. van Kampen, *Elimination of fast variables*,
Phys. Rep. 124 **70**, pp. 69-160, (1985).

[72] J. Dalibard and C. Cohen-Tannoudji, *Atomic motion in laser light: connection between semiclassical and quantum descriptions*,
J. Phys. B: At. Mol. Phys. **18**, pp. 1661-1683, (1985).

[73] I.M. Held, M. Winton, K. Takahashi, T. Delworth, F. Zeng and G. Vallis, *Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing*,
Journal of Climate **23**, pp. 2418-2427, (2010).

[74] J. Gunawardena, *A linear framework for time-scale separation in nonlinear biochemical systems*,
PLoS ONE **7**, pp. e36321, (2012).

[75] L. Michaelis L and M. Menten,*Die kinetik der Invertinwirkung*,
Biochem Z **49**, pp. 333–369,(1913).

[76] M. Santillán and H. Qian, *Irreversible thermodynamics in multiscale stochastic dynamical systems*,
Phys. Rev. E **83**, pp. 041130, (2011).

[77] C. Jarzynski, *Nonequilibrium Equality for Free Energy Differences*,
Phys. Rev. Lett. **78**, pp. 2690, (1997).

[78] C. Jarzynski, *Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach*,
Phys. Rev. E **56**, pp. 5018, (1997).

[79] G. Crooks, *Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences*,
Phys. Rev. E **60**, pp 2721, (1999).

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

[81] F. M. Ytreberg, R. H. Swendsen and D. M. Zuckerman, *Comparison of free energy methods for molecular systems*,
J. Chem. Phys. **125**, pp. 184114, (2006).

[82] W. Lechner, H. Oberhofer and C. Dellago, *Equilibrium free energies from fastswitching trajectories with large time steps*,
J. Chem. Phys. **124**, pp. 044113, (2006).

[83] H. Ge and H. Qian, *The Physical Origins of Entropy Production, Free Energy Dissipation and Their Mathematical Representations* ,
Phys. Rev. E **81**, pp. 051133, (2010).

[84] M. Esposito and C. Van den Broeck, *Three detailed fluctuation theorems*,
Phys. Rev. Lett. **104**, pp. 090601, (2010).

[85] A.E. Allahverdyan and N.H. Martirosyan, *Free energy for non-equilibrium quasi-stationary states*,
EPL (Europhysics Letters) **117**, pp. 50004, (2017).

[86] G. Mahler, *Quantum Thermodynamic Processes*,
Pan Stanford, Singapore, (2015).

[87] I.A. Martínez, É. Roldán, L. Dinis, and R.A. Rica, *Colloidal heat engines: A review*,
Soft Matter **13**, pp. 22–36, (2017).

[88] D. Chowdhury, *Resource Letter PBM-1: Physics of biomolecular machines*, Am. J. Phys. **77**, pp. 583-594, (2009).

[89] H.E.D. Scovil and E.O. Schultz-DuBois, *Three-level masers as heat engines*, Phys. Rev. Lett. **2**, pp. 262, (1959).

J.E. Geusic *et al.*, *Quantum equivalent of the Carnot cycle*, Phys. Rev. **156**, pp. 343, (1967).

[90] V.K. Konyukhov and A.M. Prokhorov, Uspekhi Fiz. Nauk, **119**, 541 (1976).

[91] E. Geva and R. Kosloff, *Three-level quantum amplifier as a heat engine: A study in finite-time thermodynamics*, Phys. Rev. E **49**, pp. 3903, (1994).

[92] E. Boukoza and D. Tannor, *Three-level systems as amplifiers and attenuators: A thermodynamic analysis*, Phys. Rev. Lett. **98**, pp. 240601, (2007).

[93] R. Uzdin, A. Levy, and R. Kosloff, *Equivalence of quantum heat machines, and quantum-thermodynamic signatures*, Phys. Rev. X **5**, pp. 031044, (2015).

[94] G. Thomas and R.S. Johal, *Expected behavior of quantum thermodynamic machines with prior information*, Phys. Rev. E **85**, pp. 041146, (2012).

[95] A. E. Allahverdyan, K. V. Hovhannisyan, A. V. Melkikh, and S. G. Gevorkian, *Carnot cycle at finite power: Attainability of maximal efficiency*, Phys. Rev. Lett. **111**, pp. 050601, (2013).

[96] N. Linden, S. Popescu, and P Skrzypczyk, *The smallest possible heat engines*, arxiv.1010.6029.

[97] A.S.L. Malabarba, A.J. Short, and P. Kammerlander, *Clock-driven quantum thermal engines*, New J. Phys. **17**, pp. 045027, (2015).

M.F. Frenzel, D. Jennings, and T. Rudolph, *Quasi-autonomous quantum thermal machines and quantum to classical energy flow*, New J. Phys. **18**, pp. 023037, (2016).

[98] M.O. Scully, *Quantum photocell: Using quantum coherence to reduce radiative recombination and increase efficiency*, Phys. Rev. Lett. **104**, pp. 207701, (2010).
M.O. Scully *et al.*, *Quantum heat engine power can be increased by noise-induced coherence*, PNAS, **108**, pp. 15097-15100, (2011).

[99] O. Abah *et al.*, *Single-ion heat engine at maximum power*,
Phys. Rev. Lett. **109**, pp. 203006, (2012).

[100] K. Zhang, F. Bariani, and P. Meystre, *Quantum optomechanical heat engine*,
Phys. Rev. Lett. **112**, pp. 150602, (2014).

[101] P.D. Mitcheson, E.M. Yeatman, G.K. Rao, A.S. Holmes, and T.C. Green, *Energy harvesting from human and machine motion for wireless electronic devices*,
Proc. IEEE **96**, pp. 1454–1486, (2008).

[102] T. Huesgen, J. Ruhhammer, G. Biancuzzi and P. Woias, *Detailed study of a micro heat engine for thermal energy harvesting*,
J. Micromech. Microeng. **20**, pp. 104004, (2010).

[103] G. J. Snyder, *Small Thermoelectric Generators*,
Electrochem. Soc. Interface **17**, pp. 54, (2008).

[104] D. Champier, *Thermoelectric generators: A review of applications*,
Energy Conversion and Management **140**, pp. 167-181, (2017).

[105] M. Smoluchowski, *Experimentell nachweisbare, der üblichen Thermodynamik widersprechende Molekularphänomene*,
Phys. Z. **13**, pp. 1069-1080, (1912).

[106] R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics*,
Addison Wesley, Reading MA **1**, chapter 46, (1963).

[107] D. Abbott, B.R. Davis and J.M.R. Parrondo, *The Problem of Detailed Balance for the Feynman-Smoluchowski Engine (FSE) and the Multiple Pawl Paradox*,
AIP Conf. Proc. **511**, pp. 213-218, (2000).

[108] P. Reimann, *Brownian motors: noisy transport far from equilibrium*,
Phys. Rep. **361**, pp. 57-265, (2002).

[109] P. Hänggi and F. Marchesoni, *Artificial Brownian motors: Controlling transport on the nanoscale*,
Rev. Mod. Phys. **81**, pp. 387, (2009).

[110] P. Hänggi, F. Marchesoni, and F. Nori, *Brownian motors*,
Ann. Phys. **14**, pp. 51-70, (2005).

[111] R. D. Astumian, *Thermodynamics and kinetics of a Brownian motor*,
Science **276**, pp. 917-922, (1997).

[112] J. M. R. Parrondo and B. J. De Cisneros, *Energetics of Brownian motors: a review*,
Applied Physics A **75**, pp. 179, (2002).

[113] I.A. Martínez, É. Roldán, L. Dinis, D. Petrov, J.M.R. Parrondo, and R.A. Rica, *Brownian carnot engine*,
Nature Physics **12**, pp. 67-70, (2016).

[114] V. Blickle and C. Bechinger, *Realization of a micrometre-sized stochastic heat engine*,
Nature Physics **8**, pp. 143-146, (2012).

[115] I.A. Martínez, É. Roldán, L. Dinis, D. Petrov, and R.A. Rica, Phys. Rev. Lett. **114**, pp. 120601, (2015).

[116] P. A. Quinto-Su, *A microscopic steam engine implemented in an optical tweezer*,
Nat. Commun. **5**, pp. 5889, (2014).

[117] S. Krishnamurthy, S. Ghosh, D. Chatterji, R. Ganapathy, and A.K. Sood, *A micrometre-sized heat engine operating between bacterial reservoirs*,
Nature Physics **12**, pp. 1134–1138, (2016).

[118] A. Ashkin, *Acceleration and trapping of particles by radiation pressure*,
Phys. Rev. Lett. **24**, pp. 156, (1970).

[119] A. Ashkin, J. Dziedzic, J. Bjorkholm and S. Chu, *Observation of a single-beam gradient force optical trap for dielectric particles*,
Opt. Lett. **11**, pp. 288–290, (1986).

[120] A.E. Allahverdyan, S.G. Babajanyan, N.H. Martirosyan, and A.V. Melkikh, *Adaptive heat engine*,
Phys. Rev. Lett. **117**, pp. 030601, (2016).

[121] S. Falk *et al.*, *Photosynthetic adjustment to temperature*, in *Photosynthesis and the environment*, pp. 367-385, ed. by N.R. Baker,
Springer, Netherlands, (1996).

[122] I. Rojdestvenski, M. G. Cottam, G. Oquist, and N. Huner, *Thermodynamics of complexity*,
Physica A **320**, pp. 318-328, (2003).

[123] R.J. Baxter, *Exactly solved models in statistical mechanics*,
Elsevier, (1982).

[124] S. Weinzierl, *Introduction to Monte Carlo methods*,
arXiv preprint hep-ph/0006269, (2000).

[125] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*,
Cambridge University Press, (1992).

[126] F. James, *Monte Carlo theory and practice*,
Rep. Prog. Phys. **43**, pp. 1145, (1980).

[127] , M.E.J. Newman and G.T.Barkema, *Monte Carlo Methods in Statistical Physics chapter 1-4*,
Oxford University Press, (1999).

[128] D.P. Kroese, T. Brereton, T. Taimre, and Z.I. Botev, *Why the Monte Carlo method is so important today*,
WIREs Comput Stat. **6**, pp. 386–392, (2014).

[129] J. Gentle, *Random Number Generation and Monte Carlo Methods*,
Springer Science & Business Media, (2006).

[130] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H.Teller, and E. Teller,*Equation of State Calculations by Fast Computing Machines*,
J Chem Phys. **21**, pp. 1087–1093, (1953).

[131] W.R. Gilks, S. Richardson and D. Spiegelhalter, *Markov chain Monte Carlo in practice*,
CRC press, (1995).

[132] G.O. Roberts and J.S. Rosenthal, *General state space Markov chains and MCMC algorithms*,
Probability Surveys **1**, pp. 20-71, (2004).

[133] I. Beichl and F. Sullivan, *The metropolis algorithm*,
Computing in Science & Engineering **2**, pp. 65-69, (2000).

[134] C. Andrieu, N. de Freitas, A. Doucet, and M. Jordan, *An introduction to MCMC for machine learning*,
Machine Learning **50**, pp. 5-43, (2003).

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

[136] G. Savvidy and N. Savvidy, *On the Monte Carlo Simulation of Physical Systems*,
J.Comput.Phys. **97**, pp. 566-572, (1991).

[137] N. Akopov, G. Savvidy and N. Savvidy, *Matrix Generator of Pseudo-Random Numbers*,
J.Comput.Phys. **97**, pp. 573-579, (1991).

[138] K. Savvidy and G. Savvidy, *Spectrum and entropy of C-systems MIXMAX random number generator*,
Chaos, Solitons & Fractals, **91**, pp. 33-38, (2016).

[139] A. Görlich, M. Kalomenopoulos, K. Savvidy and G. Savvidy, *Distribution of periodic trajectories of Anosov C-system*,
arXiv:1608.03496, (2016).

[140] S. Agostinelli et al., *Geant4: a simulation toolkit*,
Nucl. Instrum. Meth. A. **506**, pp. 250, (2003).

[141] Narek H. Martirosyan, Gevorg A. Karyan and Norayr Z. Akopov, *Statistical tests for MIXMAX pseudorandom number generator*,
Mathematical Problems of Computer Science **47**, pp. 37-49, (2017).

[142] R.L. Stratonovich, *Nonlinear Nonequilibrium Thermodynamics I*,
Springer-Verlag, Berlin, (1992).

[143] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*,
Oxford University Press, (2002).

[144] J. D. Chodera and F. Noé, *Markov state models of biomolecular conformational dynamics*,
Current opinion in structural biology **25**, pp. 135-144, (2014).

[145] C.F. Stevens, *Interactions between intrinsic membrane protein and electric field. An approach to studying nerve excitability*,
Biophysical journal **22**, pp. 295-306, (1978).

[146] B. Sakmann and E. Neher, *Single-channel recording.*
Springer Science & Business Media, New York, (2013).

[147] L. J.S. Allen, *An Introduction to Stochastic Processes with Applications to Biology*,
CRC Press, Lubbock, Texas, (2010).

[148] E. A. Codling, M. J. Plank and S. Benhamou, *Random walk models in biology*,
J. R. Soc. Interface **5**, pp. 813-834, (2008).

[149] M. Grifoni and P. Hanggi, *Driven quantum tunneling*,
Phys. Rep. **304**, pp. 229-354, (1998).

[150] W. Kohn, *Periodic thermodynamics*,
J. Stat. Phys. **103**, pp. 417-423, (2001).

[151] D.W. Hone, R. Ketzmerick and W. Kohn, *Statistical mechanics of Floquet systems: the pervasive problem of near degeneracies*,
Phys. Rev. E **79**, pp. 051129, (2009).

[152] W. Wustmann, *Statistical mechanics of time-periodic quantum systems*,
Ph.D. thesis, Technische Universität Dresden, (2010).

[153] O. Raz, Y. Subasi, C. Jarzynski, *Mimicking Nonequilibrium Steady States with Time-Periodic Driving*,
Phys. Rev. X **6**, pp. 021022, (2016) .

[154] K. Brandner, K. Saito, and U. Seifert, *Thermodynamics of micro-and nano-systems driven by periodic temperature variations*,
Phys. Rev. X **5**, pp. 031019, (2015).

[155] K. Proesmans and C. Van den Broeck, *Onsager coefficients in periodically driven systems*,
Phys. Rev. Lett. **115**, pp. 090601, (2015).

[156] N. Agmon and J.J. Hopfield, *Transient kinetics of chemical reactions with bounded diffusion perpendicular to the reaction coordinate: intramolecular processes with slow conformational changes*,
J. Chem. Phys. **78**, pp. 6947-6959, (1983).

[157] P.L. Bhatnager, E.P. Gross and M. Krook, *A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems*,
Phys. Rev. **94**, pp. 511, (1954).

[158] M. Krook, *Dynamics of rarefied gases*,
Phys. Rev. **99**, pp. 1896, (1955).

[159] E.P. Gross and M. Krook, *Model for collision processes in gases: Small-amplitude oscillations of charged two-component systems*,
Phys. Rev. **102**, pp. 593, (1956).

[160] E.P. Gross, E.A. Jackson and S. Ziering, *Boundary value problems in kinetic theory of gases*,
Ann. Phys. **1**, pp. 141-167, (1957).

[161] E.P. Gross and E.A. Jackson, *Boundary value problems in kinetic theory of gases*,
Phys. Fluids **2**, pp. 432-441, (1959).

[162] H. Grad, *Principles of the Kinetic Theory of Gases*,
in Handbuch der Physik, ed. by S. Flugge, Springer-Verlag, (1958).

[163] C. De Dominicis, H. Orland and F. Lainée, *Stretched exponential relaxation in systems with random free energies*,
Journal de Physique Lettres **46**, pp. 463-466, (1985).

[164] S. Denisov, S. Flach and P Hanggi, *Tunable transport with broken space–time symmetries*,
Phys. Rep. **538**, pp. 77-120, (2014).

[165] D. N. Zubarev, V. Morozov, G. Ropke, *Statistical Mechanics of Nonequilibrium Processes.*
John Wiley & Sons, NY, (1996).

[166] J. Meixner, in *A Critical Review of Thermodynamics*, edited by E.B. Stuart, B. Gal-Or, and A. J. Brainard (Mono Book Corporation, Baltimore, MD, 1970), p. 40.

[167] R. Balian, *From Microphysics to Macrophysics*,
volume I, Springer, (1992).

[168] C. Jarzynski, *Thermalization of a Brownian particle via coupling to low-dimensional chaos*,
Phys. Rev. Lett. **74**, pp. 2937, (1995).

[169] A. Riegert, N. Baba, K. Gelfert, W. Just, and H. Kantz, *Hamiltonian chaos acts like a finite energy reservoir: Accuracy of the Fokker-Planck approximation*,
Phys. Rev. Lett. **94**, pp. 054103, (2005).

[170] K. Sekimoto, *Stochastic Energetics*,
Springer, Berlin, (2010).

[171] L.F. Cugliandolo, J. Kurchan, and L. Peliti, *Energy flow, partial equilibration, and effective temperatures in systems with slow dynamics*,
Phys. Rev. E **55**, pp. 3898, (1997).

[172] J. Casas-Vazquez and D. Jou, *Temperature in non-equilibrium states: a review of open problems and current proposals*,
Rep. Prog. Phys. **66**, pp. 1937, (2003).

[173] L. F. Cugliandolo, *The effective temperature*,
J. Phys. A **44**, pp. 483001, (2011).

[174] E. Bertin, K. Martens, O. Dauchot, and M. Droz, *Intensive thermodynamic parameters in nonequilibrium systems*,
Phys. Rev. E **75**, pp. 031120, (2007).

[175] T.M. Nieuwenhuizen, *Thermodynamic picture of the glassy state gained from exactly solvable models*,
Phys. Rev. E **61**, pp. 267, (2000);
Formulation of thermodynamics for the glassy state: Configurational energy as a modest source of energy,
J. Chem. Phys. **115**, pp. 8083-8088, (2001).

[176] A. Celani, S. Bo, R. Eichhorn, and E. Aurell, *Anomalous thermodynamics at the microscale*,
Phys. Rev. Lett. **109**, pp. 260603, (2012).

[177] M. Esposito, *Stochastic thermodynamics under coarse graining*,
Phys. Rev. E **85**, pp. 041125, (2012).

[178] A. Puglisi, S. Pigolotti, L. Rondoni, and A. Vulpiani, *Entropy production and coarse graining in Markov processes*,
JSTAT, pp. P05015, (2010).

[179] A. Gomez-Marin, J.M.R. Parrondo, and C. Van den Broeck, *Lower bounds on dissipation upon coarse graining*,
Phys. Rev. E **78**, pp. 011107, (2008).

[180] T. Tom T and de M.J. Oliveira, *Stochastic approach to equilibrium and nonequilibrium thermodynamics*,
Phys. Rev. E **91**, pp. 042140, (2015).

[181] R. Ziener, A. Maritan and H. Hinrichsen, *On entropy production in nonequilibrium systems*,
JSTAT, pp. P08014, (2015).

[182] R. Landauer and J. Woo, *Entropy changes in the steady state far from equilibrium*,
Phys. Rev. A **6**, pp. 2204, (1972).

[183] A.E. Allahverdyan, Th.M. Nieuwenhuizen, *Steady adiabatic state: Its thermodynamics, entropy production, energy dissipation, and violation of Onsager relations*,
Phys. Rev. E **62**, pp. 845, (2000).

[184] A.C.C. Coolen, R.W. Penney, and D. Sherrington, *Coupled dynamics of fast spins and slow interactions in neural networks and spin systems*,
J. Phys. A **26**, pp. 3681, (1993).

[185] R. Penney and D. Sherrington, *Slow interaction dynamics in spin-glass models*,
J. Phys. A **27**, pp. 4027, (1994).

[186] V. Dotsenko, S. Franz and M. Mezard, *Partial annealing and overfrustration in disordered systems*,
J. Phys. A **27**, pp. 2351, (1994).

[187] A.E. Allahverdyan and K. G. Petrosyan, *Anomalous latent heat in nonequilibrium phase transitions.*
Phys. Rev. Lett. **96**, pp. 065701, (2006).

[188] O.M. Ritter, P.C.T. D'Ajello, and W. Figueiredo, *Fast and slow degrees of freedom coupling two different reservoirs*,
Phys. Rev. E **69**, pp. 016119, (2004).

[189] P. Ilg, and J.-L. Barrat, *Effective temperatures in a simple model of non-equilibrium, non-Markovian dynamics*,
Journal of Physics: Conference Series **40**, pp. 76, (2006).

[190] I. Rojdestvenski *et al.*, *Robustness and time-scale hierarchy in biological systems*,
BioSystems **50**, pp. 71-82, (1999).

[191] J. Gunawardena, *A linear framework for time-scale separation in nonlinear biochemical systems*,
PLoS ONE **7**, pp. e36321, (2012).

[192] G.A. Pavliotis and A.M. Stuart, *Multiscale methods: averaging and homogenization*,
Springer Science+Business Media, NY, (2008).

[193] J.M.G. Vilar, and J.M. Rubi, *Work-Hamiltonian connection for anisoparametric processes in manipulated microsystems*,
J. Non-Equilib. Thermodyn. **36**, pp. 123-130, (2011).

[194] A. E. Allahverdyan, K. Hovhannisyan, D. Janzing and G. Mahler, *Thermodynamic limits of dynamic cooling*,
Phys. Rev. E **84**, 041109 (2010).

[195] A. Abragam and M. Goldman, *Principles of dynamic nuclear polarisation*,
Rep. Prog. Phys. **41**, pp. 395, (1978).

O.W. Sorensen, *Polarization transfer experiments in high-resolution NMR spectroscopy*,
Prog. Nucl. Magn. Reson. Spectrosc. **21**, 503-569, (1989).

[196] W. Ketterle and D. E. Pritchard, *Atom cooling by time-dependent potentials*,
Phys. Rev. A **46**, pp. 4051, (1992).

[197] A. Bartana, R. Kosloff, and D. J. Tannor, *Laser cooling of internal degrees of freedom. II*,
J. Chem. Phys. **106**, pp. 1435-1448, (1997).

[198] J.M. Fernandez, S. Lloyd, T. Mor, and V. Roychowdhury, *Algorithmic cooling of spins: A practicable method for increasing polarization*,
Int. J. Quantum Inform. **2**, pp. 461-477, (2004).

[199] L.-A. Wu, D. Segal, and P. Brumer, *Ground state cooling is not possible given initial system-thermal bath factorization*,
Scientific Reports **3**, pp. 1824, (2013).

[200] H.J. Briegel and S. Popescu, *Intra-molecular refrigeration in enzymes*,
Proc. R. Soc. A **469**, pp. 20110290, (2013).

[201] E.B. Mpemba and D.G. Osborne, *The Mpemba effect*,
Physics Education, **4**, pp. 172-175, (1969).

[202] G.S. Kell, *The freezing of hot and cold water*,
Am. J. Phys. **37**, pp. 564, (1969).

[203] D. Auerbach, *Supercooling and the Mpemba effect: When hot water freezes quicker than cold*,
Am. J. Phys. **63**, pp. 882-885, (1995).

[204] H.C. Burridge and P.F. Linden, *Questioning the Mpemba effect: hot water does not cool more quickly than cold*,
Scientific Reports **6**, pp. 37665, (2016).

[205] Z. Lu and O. Raz, *Anomalous cooling and heating-the Mpemba effect and its inverse*, arXiv:1609.05271.

[206] J.I. Katz, *When hot water freezes before cold*,
Am. J. Phys. **77**, pp. 27-29, (2009).

[207] H.B. Callen, *Thermodynamics and an introduction to thermostatistics*,
John Wiley & Sons, NY, (1985).

[208] S. Kauffman, *Molecular autonomous agents*,
Phil. Trans. R. Soc. A **361**, pp. 1089-1099, (2003).

[209] D. La Manna *et al.*, *Reconfigurable electrical interconnection strategies for photovoltaic arrays: A review*,
Renewable and Sustainable Energy Reviews, **33**, pp. 412-426, (2014).

[210] E. Albarran-Zavala and F. Angulo-Brown, *A simple thermodynamic analysis of photosynthesis*,
Entropy, **9**, pp. 152-168, (2007).

[211] T. Friedlander and N. Brenner, *Adaptive response by state-dependent inactivation*,
PNAS **106**, pp. 22558-22563, (2009).

[212] M. Inoue and K. Kaneko, *Dynamics of coupled adaptive elements: Bursting and intermittent oscillations generated by frustration in networks*,
Phys. Rev. E **81**, pp. 026203, (2010).

[213] G. Lan, P. Sartori, S. Neumann, V. Sourjik, Y. Tu, *The energy-speed-accuracy trade-off in sensory adaptation*,
Nature Physics **8**. pp. 422-428, (2012).

[214] A. E. Allahverdyan and Q.A. Wang, *Adaptive machine and its thermodynamic costs*,
Phys. Rev. E **87**, 032139 (2013).

[215] A. C. Barato, D. Hartich, and U. Seifert, *Information-theoretic versus thermodynamic entropy production in autonomous sensory networks*,
Phys. Rev. E **87**, pp. 042104, (2013). *Efficiency of cellular information processing*,
New J. Phys. **16**, 103024 (2014).

[216] P. Sartori and Y. Tu, *Free energy cost of reducing noise while maintaining a high sensitivity*,
Phys. Rev. Lett. **115**, pp. 118102, (2015).

[217] D. Markovic and C. Gros, *Self-organized chaos through polyhomeostatic optimization*,
Phys. Rev. Let. **105**, pp. 068702, (2010).

[218] S. McGregor and N. Virgo, in *Advances in Artificial Life: Darwin Meets von Neumann*, pp. 230-237,
Springer, Berlin, (2009).

[219] K. Proesmans, B. Cleuren, and C. Van den Broeck, , *Power-efficiency-dissipation relations in linear thermodynamics*,
Phys. Rev. Lett. **116**, pp. 220601, (2016).

[220] R. Zwanzig, *Rate processes with dynamical disorder*,
Acc. Chem. Res. **23**, pp. 148-152, (1990).

[221] H. Frauenfelder, P.G. Wolynes, and R.H. Austin, *Biological physics*,
Rev. Mod. Phys. **71**, S419 (1999).

[222] M. Kurzynski, *The thermodynamic machinery of life*,
Springer-Verlag, Berlin, (2010).

[223] L.A. Blumenfeld and A.N. Tikhonov, *Biophysical Thermodynamics of Intracellular Processes*,
Springer, Berlin, (1994).

[224] H.-P. Lerch, A.S. Mikhailov, and B. Hess, *Conformational-relaxation models of single-enzyme kinetics*,
PNAS, **99**, pp. 15410-15415, (2002). H.-P. Lerch, R. Rigler, A.S. Mikhailov, *Functional conformational motions in the turnover cycle of cholesterol oxidase*,
PNAS **102**, pp. 10807-10812, (2005).

[225] C. Bustamante, Y.R. Chemla, N.R. Forde, and D. Izhaky, *Mechanical processes in biochemistry*,
Annu. Rev. Biochem. **73**, pp. 705-748, (2004).

Z.D. Nagel and J.P. Klinman, *A 21st century revisionist's view at a turning point in enzymology*,
Nat. Chem. Biol. **5**, pp. 543-550, (2009).

[226] K.V. Shaitan and A.B. Rubin, *Stochastic dynamics and electron-conformation interactions in proteins*,
Biofizika, **30**, pp. 517-526, (1984). K.V. Shaitan, *Dynamics of electron-conformational transitions and new approaches to the physical mechanisms of functioning of biomacromolecules*,
Biophysics, **39**, pp. 993-1012, (1994)

[227] A. L. Borovinskiy and A. Yu. Grosberg, *Design of toy proteins capable of rearranging conformations in a mechanical fashion*,
J. Chem. Phys. **118**, pp. 5201-5212, (2003).

[228] V.A. Kovarskii, *Quantum processes in biological molecules. Enzyme catalysis*,
Physics-Uspekhi **42**, pp. 797, (1999).

[229] A.O. Gouschcha *et al.*, *Self-regulation phenomena in bacterial reaction centers. I. General theory*,
Biophys. J. **79**, pp. 1237-1252, (2000).

L.N. Christophorov, A.R. Holzwarth, V.N. Kharkyanen, and F. van Mourik, *Structure–function self-organization in nonequilibrium macromolecular systems*,
Chemical Physics **256**, pp. 45-60, (2000)

[230] A.I. Berg *et al.*, *CONFORMATIONAL REGULATION OF ACTIVITY IN PHOTOSYNTHETIC MEMBRANES OF PURPLE BACTERIA*,
Molecular Biology **13**, pp. 58-74, (1979).

[231] Yu. I. Samoilenko, *NEGATIVE SENSITIVITY OF DYNAMIC-SYSTEMS AND ITS UTILIZATION FOR STABILIZATION OF UNSTABLE PROCESSES*,
Autom. Remote Control **39**, pp. 632-640, (1978); *ibid.* **39**, 1745 (1979).

[232] See https://en.wikipedia.org/wiki/Negative resistance

[233] V.P. Starr, *Physics of negative viscosity phenomena*,
McGraw-Hill, (1968).

[234] B. Cleuren and C. Van den Broeck, *Random walks with absolute negative mobility*,
Phys. Rev. E. **65**, pp. 030101, (2002). A. Haljas, R. Mankin, A. Sauga, E. Reiter, *Anomalous mobility of Brownian particles in a tilted symmetric sawtooth potential*,
Phys. Rev. E. **70**, pp. 041107. (2004). J. Spiechowicz, J. Luczka, P. Hanggi, *Absolute negative mobility induced by white Poissonian noise*,
JSTAT **2013**, pp. P02044, (2013).

[235] A.E. Allahverdyan and Th.M. Nieuwenhuizen, *Steady adiabatic state: Its thermodynamics, entropy production, energy dissipation, and violation of Onsager relations*,
Phys. Rev. E **62**, pp. 845, (2000).

[236] N.F. Ramsey, *Thermodynamics and statistical mechanics at negative absolute temperatures*,
Phys. Rev. **103**, pp. 20, (1956); S. Machlup, *Negative temperatures and negative dissipation*,
Am. J. Phys. **43**, pp. 991-995, (1975).

[237] A.G. Butkovskii and Yu.I. Samoilenko, *Control of Quantum Mechanical Processes*,
Springer, Berlin, (1991).

[238] H. Schmidt and G. Mahler, *Control of local relaxation behavior in closed bipartite quantum systems*,
Phys. Rev. E **72**, pp. 016117, (2005).

[239] N. Brunner, N. Linden, S. Popescu, and P. Skrzypczyk, *Virtual qubits, virtual temperatures, and the foundations of thermodynamics*,
Phys. Rev. E **85**, pp. 05111 (2012).

[240] S. Braun, P. Ronzheimer, M. Schreiber, S. S. Hodgman, T. Rom, I. Bloch, and U. Schneider, *Negative absolute temperature for motional degrees of freedom*,
Science, **339**, pp. 52-55, (2013).

[241] J. Dongarra and F. Sullivan, *Guest editors' introduction: The top 10 algorithms*,
Computing in Science & Engineering **2**, pp. 22-23, (2000).

[242] G. Marsaglia, *DIEHARD: a battery of tests of randomness*,
http://stat.fsu.edu/pub/diehard/, (1996).

[243] M. Mascagni and A. Srinivasan, *Algorithm 806: SPRNG: A scalable library for pseudorandom number generation*,
TOMS **26**, pp. 436-461, (2000).

[244] A. Rukhin, et al., *A statistical test suite for random and pseudorandom number generators for cryptographic applications*,
DTIC Document, (2001).

[245] P. L'ecuyer and R. Simard, *TestU01: A C Library for Empirical Testing of Random Number Generators*,
ACM Transactions on Mathematical Software **33**, pp. 22, (2007).

[246] D. Knuth, *The Art of Computer Programming, Seminumerical Algorithms, Volume 2, 3rd edition*,
Massachusetts: Addison Wesley, (1998).

[247] R.J. Serfling, *Approximation theorems of mathematical statistics*,
New York: John Wiley & Sons, (2009).

[248] T.B. Arnold and J.W. Emerson, *Nonparametric goodness-of-fit tests for discrete null distributions*,
The R Journal **3**, pp. 34-39, (2011).

[249] T.W. Anderson and D.A. Darling, *Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes*,
The Annals of Mathematical Statistics, pp. 193-212, (1952).

[250] J. Wang, W. W. Tsang, and G. Marsaglia, *Evaluating Kolmogorov's distribution*,
Journal of Statistical Software **8**, (2003).

[251] J. A. Peacock, *Two-dimensional goodness-of-fit testing in astronomy*,
Monthly Notices Royal Astronomy Society **202**, pp. 615-627, (1983).

[252] G. Fasano and A. Franceschini, *A multidimensional version of the Kolmogorov-Smirnov test*,
Monthly Notices Royal Astronomy Society **225**, pp. 155-170, (1987).

[253] N.Z. Akopov and N.H. Martirosyan, *The Optimal Approach for Kolmogorov-Smirnov Test Calculation in High Dimensional Space*,
Mathematical Problems of Computer Science **44**, pp. 138-144, (2015).

[254] A. Justel, D. Pena, R. Zamar, *A multivariate Kolmogorov-Smirnov test of goodness of fit*,
Statistics & Probability Letters **35**, pp. 251-259, (1997).

[255] E. Gosset, *A three-dimensional extended Kolmogorov-Smirnov test as a useful tool in astronomy*,
Astronomy and Astrophysics **188**, pp. 258-264, (1987).

[256] D.J. Hudson, *Lectures on elementary statistics and probability*,
CERN: European Organization for Nuclear Research, (1964).

[257] S. Tezuka,*Uniform random numbers: Theory and practice*,
Springer Science & Business Media, (2012).

[258] P. L'ecuyer, *Efficient and portable combined random number generators*,
Commun. ACM **31**, pp. 742-751, (1988).

[259] P. L'ecuyer, *Tables of linear congruential generators of different sizes and good lattice structure*,
Mathematics of Computation **68**, pp. 249-260, (1999).

[260] U. Dieter, *How to calculate shortest vectors in a lattice*,
Mathematics of Computation **29**, pp. 827-833, (1975).

[261] G. Fishman, *Monte Carlo: concepts, algorithms, and applications*,
Springer Science & Business Media, (2013).

[262] , P. Hellekalek, *Good random number generators are (not so) easy to find*,
Mathematics and Computers in Simulation **46**, pp. 485-505, (1998).

[263] M. Mascagni, *Parallel pseudorandom number generation*,
SIAM News **32**, pp. 221-251, (1999).

[264] A. Srinivasan, D.M. Ceperley, and M. Mascagni, *Random number generators for parallel applications*,
Advances in chemical physics **105**, pp. 13-36, (1999).

[265] P. Frederickson, R. Hiromoto, T.L. Jordan, B. Smith and T. Warnock, *Pseudo-random trees in Monte Carlo*,
Parallel Computing **1**, pp. 175-150, (1984).

[266] D. Istvan, *Uniform random number generators for parallel computers*,
Parallel Computing **15**, pp. 155-164, (1990).

[267] A. Srinivasan, M. Mascagni and D. Ceperley, *Testing parallel random number generators*,
Parallel Computing **29**, pp. 69-94, (2003).

[268] http://cernlib.web.cern.ch/cernlib/

[269] T. Bradley, J. du Toit, R. Tong, M. Giles and P. Woodhams, *Parallelization techniques for random number generations*,
GPU Computing Gems Emerald Edition **16**, pp. 231-246, (2011).

[270] J.K. Salmon, M.A. Moraes, R.O. Dror and D.E. Shaw, *Parallel random numbers: as easy as 1, 2, 3*,
Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis (SC), New York, pp. 1-12, (2011).

[271] D.R.C. Hill, C. Mazel, J. Passerat-Palmbach and M.K. Traore, *Distribution of random streams for simulation practitioners*,
Concurrency and Computation: Practice and Experience **25**, pp. 1427-1442, (2013).

[272] N.L. Johnson, S. Kotz and N. Balakrishnan, *Continuous univariate distributions*,
Wiley series in probability and mathematical statistics, New York, (1995).