Phase transitions in thermodynamically consistent biochemical systems

Von der Fakultät Mathematik und Physik der Universität Stuttgart
zur Erlangung der Würde eines Doktors der
Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

vorgelegt von

Basile Nguyen

aus Lausanne (Schweiz)

Hauptberichter: Prof. Dr. Udo Seifert
Mitberichter: Prof. Dr. Siegfried Dietrich
Vorsitzender: Prof. Dr. Martin Dressel

Tag der Einreichung: 05. Juni 2020
Tag der mündlichen Prüfung: 22. Juli 2020

II. Institut für Theoretische Physik der Universität Stuttgart

2020

Contents

List of Figures

Publications

Parts of this thesis have been published in

- B. Nguyen and U. Seifert,
  Exponential volume dependence of entropy-current fluctuations at first-order phase transitions in chemical reaction networks,
  Phys. Rev. E **102**, 022101 (2020).

   Chapter 3 is based on this publication.

- B. Nguyen, U. Seifert and A.C. Barato,

  Phase transition in thermodynamically consistent biochemical oscillators,
  J. Chem. Phys. **149**, 045101 (2018).

   Chapter 4 is based on this publication.

Further publications I have co-authored

- J. H. Fritz, B. Nguyen and U. Seifert,
  Stochastic thermodynamics of chemical reactions coupled to finite reservoirs: A case study for the Brusselator,
  J. Chem. Phys. **152**, 235101 (2020).

- A. Ehrmann, B. Nguyen, U. Seifert,
  Interlinked GTPase cascades provide a motif for both robust switches and oscillators,
  J. R. Soc. Interface 16, 20190198 (2019).

- B. Nguyen, D. Hartich, U. Seifert and P. De Los Rios,
  Thermodynamic bounds on the ultra- and infra-affinity of Hsp70 for its substrates,
  Biophys. J. **113**, 362-370 (2017).

Zusammenfassung

Lebende Systeme benötigen zur Erfüllung ihrer Aufgaben eine konstante Energiezufuhr. Viele biologische Systeme benötigen Energieträger, die sogenannten Nukleotidphosphate, wie z.B. Adenosintriphosphat (ATP). Chemische Energie wird bei einer Hydrolysereaktion durch Aufbrechen einer der Phosphatbindungen erzeugt. Diese Reaktion verwandelt ein ATP in ein Adenosindiphosphat (ADP) und ein Phosphat (P_i). Die Zellen halten einen ATP-Exzess durch einen metabolischen Prozess, die sogenannte Zellatmung, aufrecht, wobei ADP kontinuierlich in ATP umwandelt wird. Zur Vereinfachung können wir davon ausgehen, dass ein biologisches System an chemische Reservoirs gekoppelt ist, die viel größer sind als das System und dass sich diese Reservoirs auf den relevanten Zeitskalen daher nicht signifikant ändern. Wenn das System nicht von außen gesteuert wird, kommt es ab einem gewissen Zeitpunkt zu einem stationären Nicht-Gleichgewichtszustand.

Biologische Systeme beruhen auf Bistabilität und Oszillationen, um entscheidende Prozesse zu regulieren. Einerseits steuern Schalter die zelluläre Reaktion auf ein Input-Signal über komplexe Rückkopplungsstrukturen und sind zentrale Elemente vieler zellulärer Prozesse wie der Genexpression, des Membrantransports oder des Vesikeltransports. Ein idealer Schalter sollte ein all-or-none Response aufweisen und auf Bistabilität beruhen, um gegenüber Input-Störungen robust zu sein. Andererseits steuern biochemische Oszillatoren das Timing und die Kontrolle der zentralen Prozesse der Zelle wie zirkadiane Rhythmen und den Zellzyklus. Diese zwei Regime sind mit Phasenübergängen erster und zweiter Ordnung verbunden. In dieser Dissertation werden wir den Phasenübergänge charakterisieren, der in biochemischen Systemen stattfindet.

Laut der Thermodynamik können komplexe Verhaltensweisen in chemischen Reaktionsnetzwerken wie Bistabilität oder Oszillationen nur weit außerhalb des Gleichgewichts erreicht werden. Im Gleichgewicht wird das System in einen eindeutigen Gleichgewichtszustand relaxieren. Das Verhalten des Systems hängt von seinem Abstand vom Gleichgewicht ab und muss an einem gewissen Punkt zu einem Phasenübergang kommen, um Oszillationen auszubilden. Bei Systemen, die in Kontakt mit chemischen Reservoirs sind, kann die Differenz der chemischen Potentiale zwischen den Reservoirs ein Steuerparameter für diesen Phasenübergang sein. Diese Differenz führt zu thermodynamischen Kräften, die das System in einem Nicht-Gleichgewichtszustand treiben. Zum Beispiel hängt in Zellen die thermodynamische Kraft von der Differenz der chemischen Potentiale zwischen ATP, ADP und P_i ab.

Nichtgleichgewichts-Phasenübergänge wurden seit langem untersucht und sind immer noch weniger verstanden als ihre Pendants im Gleichgewicht. Das Konzept der Ordnungsparameter kann auf ähnliche Weise wie bei Gleichgewichtssystemen verwendet werden. Allerdings kann es nicht den irreversiblen Charakter eines Nichtgleichgewichts-Phasenübergangs beschreiben. Die neuesten Fortschritte in der stochastischen Thermodynamik haben die Wichtigkeit der Entropieproduktion als Quantifikator der Entfernung vom Gleichgewicht betont, da sie in einem Gleichgewichtssystem verschwindet und in einem Nicht-Gleichgewichtsregime positiv ist. Daraus folgt die Frage, ob die Entropieproduktion ein charakteristisches Verhalten bei Phasenübergängen zeigt.

Die Erforschung von Stromfluktuationen hat sich seit der Entdeckung der "Thermodynamischen Unschärferelation" im Jahr 2015 wesentlich weiterentwickelt. Diese Relation besagt, dass die Präzision jedes Stroms einen minimalen thermodynamischen Kostenfaktor hat. Sie gilt für eine breite Palette von Systemen, die durch einen Markov-Prozess beschrieben werden. Ein besonders wichtiger Strom ist die totale Entropieproduktion. Obwohl die Entropie-Stromfluktuationen eine universelle untere Schranke haben, sind sie von oben nicht beschränkt, insbesondere in der Nähe eines Phasenübergangs, bei dem wir erwarten, dass die Fluktuationen thermodynamischer Größen divergieren. Das Hauptziel dieser Dissertation ist eine Charakterisierung von Nichtgleichgewichts-Phasenübergängen. Insbesondere interessieren wir uns für das Verhalten der Entropieproduktion bei Phasenübergängen erster und zweiter Ordnung.

Kapitel 2: Grundlagen

Dieses Kapitel bietet einen Überblick über die Grundlagen der stochastischen Dynamik, der stochastischen Thermodynamik und der chemischen Reaktionsnetzwerke. Ausgehend von einer grundlegenden Theorie der statistischen Mechanik beginnen wir mit einem geschlossenen System in Kontakt mit einem Wärmebad und identifizieren Mesozustände. Angenommen das System äquilibriert lokal innerhalb jedes Mesozustände, dann wird die Dynamik durch einen Markov-Prozess auf der mesoskopischen Ebene beschrieben. Ein geschlossenes System wird ein thermisches Gleichgewicht erreichen, daher müssen die Übergangsraten zwischen den Mesozuständen eine detaillierte Gleichgewichtsbedingung erfüllen, die die freien Energien der Mesozustände beinhaltet. Durch die Aufteilung eines geschlossenen Systems in ein Kernsystem und chemische Reservoirs können wir die Dynamik offener Systeme beschreiben. Durch entsprechende Steuerung der chemischen Potenziale der Reservoirs kann ein offenes System im Gegensatz zu seinem geschlossenen Pendant einen stationären Nicht-Gleichgewichtszustand erreichen. Dann identifizieren wir thermodynamische Observablen wie Wärme-, Arbeits- oder Entropieproduktion entlang einzelner Trajektorien so, dass sich die erwarteten makroskopischen Observablen ergeben.

Daraufhin betrachten wir die Dynamik chemischer Reaktionsnetzwerke. Mit der An-

nahme, dass chemische Reaktionen Poisson-Prozesse sind, definieren wir Übergangsraten und führen die chemische Master-Gleichung (CME) ein, die die Dynamik eines chemischen Reaktionsnetzwerks beschreibt. Es ist im Allgemeinen sehr schwierig diese Master-Gleichung exakt zu lösen. Wir stellen daher den stochastischen Simulationsalgorithmus vor, der eine exakte Methode zur Erzeugung stochastischer Trajektorien liefert. Zunächst leiten wir die chemische Langevin-Gleichung ("CLE") aus der CME ab, indem wir die Poisson-Zufallsvariablen mit normalverteilten Zufallsvariablen annähern. Die CLE kann auch als Fokker-Planck-Gleichung formuliert werden, die man normalerweise durch Trunkierung der chemischen Kramers-Moyal-Gleichung erhält. Wir leiten die lineare Rauschnäherung nach van Kampen ab, indem wir die CLE linearisieren. Im Limes unendlicher Systemgröße erhalten wir die deterministischen Ratengleichungen. Schließlich diskutieren wir, ob trimolekulare Reaktionen sinnvoll sind. Wir zeigen, dass eine trimolekulare Reaktion auf der stochastischen Ebene die gleiche Dynamik haben kann wie eine Reihe bimolekularer Reaktionen, bei denen eine zusätzliche kurzlebige Spezies eingeführt werden muss.

Kapitel 3: Phasenübergang erster Ordnung bei biochemischen Schaltern

In diesem Kapitel betrachten wir biochemische Schalter, die bei ihrer Aktivierung einen Phasenübergang erster Ordnung durchlaufen. Aus einer deterministischen Perspektive ist dieser Phasenübergang mit dem Konzept der Bistabilität verbunden, bei der zwei stabile stationäre Zustände koexistieren können. Im Gegensatz dazu ist in der stochastischen Perspektive der stationäre Zustand eindeutig und mit einer bimodalen Verteilung verbunden. Bei Phasenübergängen erster Ordnung ist die Entropieproduktionsrate bekanntermaßen diskontinuierlich. Wir zeigen, dass die Fluktuationen der Entropieproduktion in der Nähe des bistabilen Punktes eine exponentielle Volumenabhängigkeit haben. Unsere Ergebnisse, die wir für ein paradigmatisches Modell biochemischer Schalter, das Schlögl'sche Modell, erhalten haben, können auf eine breite Klasse von Nichtgleichgewichtssystemen angewendet werden.

Wir beginnen mit der chemischen Master-Gleichung. Mit der Theorie der großen Abweichungen berechnen wir die mittlere Rate der Entropieproduktion und den zugehörigen Diffusionskoeffizienten numerisch. Im Limes großer Systemgrößen gibt es zwei relevante Zeitskalen: eine schnelle Relaxation zum nächsten Fixpunkt und einen langsameren Übergang zwischen den Fixpunkten. In der Nähe der Fixpunkte kann das System durch stationäre Gauß'sche Prozesse modelliert werden. Für die langsamere Dynamik reduzieren wir die CME auf ein Zwei-Zustand Modell, in dem wir zeigen, dass die vergröberten Übergangsraten proportional zum Exponential des inversen Volumens sind. Für dieses Zwei-Zustands-Modell berechnen wir einen analytischen Ausdruck für den Diffusionskoeffizienten und zeigen, dass er eine exponentielle Volumenabhängigkeit am

bistabilen Punkt besitzt, wobei der exponentielle Vorfaktor durch die Höhe der effektiven Potentialbarriere, die die beiden Fixpunkte trennt, gegeben ist.

Die Berechnung der Diffusionskoeffizienten für ein Zwei-Zustands-Modell ist relativ einfach, die Lösung wird jedoch in Form der vergröberten Übergangsraten ausgedrückt. Die Ableitung dieser Raten beginnend mit der CME kann eine anspruchsvolle Aufgabe sein, zu sehen im Appendix dieses Kapitels. In einem alternativen Zugang beginnend mit der Fokker-Planck-Gleichung leiten wir einen analytischen Ausdruck für den Diffusionskoeffizienten her. Für bistabile Systeme beschreibt die zweifach verkürzte chemische Kramers-Moyal-Gleichung nur die Dynamik auf der kürzeren Zeitskala richtig und hat außerdem nicht die richtige stationäre Verteilung. Hier betrachten wir die Fokker-Planck-Gleichung mit einem effektiven Diffusionsterm, der zu derselben stationären Verteilung wie die CME führt und die Dynamik auf der längeren Zeitskala korrekt beschreibt.

Kapitel 4: Phasenübergang zweiter Ordnung in biochemischen Oszillatoren

Biochemische Oszillationen sind in lebenden Organismen omnipräsent. In einem autonomen System, das nicht durch ein externes Signal beeinflusst wird, können sie nur im Nichtgleichgewicht auftreten. Wir zeigen, dass sie durch einen generischen Nicht-Gleichgewichts-Phasenübergang entstehen, mit einem charakteristischen qualitativen Verhalten bei Kritikalität. Wir untersuchen das kritische Verhalten des Entropieproduktionsstroms für drei Versionen bekannter Modelle: das Brusselator-Modell, das Aktivator-Inhibitor-Modell und ein Modell für KaiC-Oszillationen. Wir stellen fest, dass die erste Ableitung der mittleren Rate der Entropieproduktion eine Diskontinuität am kritischen Punkt besitzt, an dem die Fluktuationen als Potenzgesetz mit dem Volumen divergieren.

Anschließend werden Metriken für die Präzision der biochemischen Oszillationen diskutiert. Als erstes betrachten wir das Standardmaß für die Präzision, die Anzahl der kohärenten Oszillationen, d.h. die Anzahl der Perioden, für die verschiedene stochastische Realisationen miteinander kohärent sind. Zweitens betrachten wir den Fano-Faktor, der mit dem thermodynamischen Fluss verbunden ist. Angesichts der thermodynamischen Unschärferelation ist der Fano-Faktor eine logische Wahl, um die Präzision mit der Thermodynamik in Beziehung zu setzen. Außerdem wurde er als Messwert vorgeschlagen, um die Präzision der biochemischen Oszillatoren zu quantifizieren. Da der Fano-Faktor selbst dann klein sein kann, wenn es keine biochemischen Oszillationen gibt, argumentieren wir, dass die Anzahl der kohärenten Oszillationen besser geeignet ist, um die Präzision der biochemischen Oszillationen zu quantifizieren.

Ausblick

Das Hauptziel dieser Arbeit war es, das Verhalten des Entropieproduktionsstroms und seiner Fluktuationen bei Phasenübergängen erster und zweiter Ordnung in homogenen chemischen Reaktionsnetzwerken zu beschreiben. Für chemische Reaktionsnetzwerke mit räumlicher Dynamik, in denen chemische Reaktionen lokal stattfinden und sich die chemische Spezies durch Diffusion ausbreitet, bleibt die Charakterisierung der Entropieproduktion eine offene Frage. Überraschenderweise bleibt die gesamte Entropieproduktion in solchen Systemen beim Phasenübergang analytisch.

Wir haben die Präzision der biochemischen Oszillationen in autonomen Systemen untersucht. Eine interessante zukünftige Forschungsrichtung wären biochemische Oszillatoren, die von einem externen periodischen Signal angetrieben werden. Für solche Systeme wurde gezeigt, dass die Präzision der Oszillationen unendlich sein kann. Außerdem hat sich in letzter Zeit gezeigt, dass Stromfluktuationen verschwinden können und damit die ursprüngliche thermodynamische Unschärferelation verletzt wird. Eine weitere unerwartete Möglichkeit, die Präzision der Oszillationen zu verbessern, ist die Berücksichtigung endlicher chemischer Reservoirs. Wir haben kürzlich gezeigt, dass sie ihre idealen Pendants übertreffen können.

Summary

Living systems require a constant supply of energy to perform their tasks and they must operate far from equilibrium. Most biological systems rely on energy carriers called nucleotide phosphates such as adenosine triphosphate (ATP). Chemical energy can be extracted through a hydrolysis reaction by breaking one of the phosphate bonds. This reaction converts an ATP into an adenosine diphosphate (ADP) and an inorganic phosphate (P_i). Cells maintain an excess of ATP through a metabolic process called cellular respiration that continuously recycles ADP into ATP. For simplicity, we can assume that a biological system is coupled to chemical reservoirs that are much larger than the system and thus these reservoirs do not evolve in a significant way on relevant timescales. If the system is not externally controlled, it will at some point in time reach a stationary state, which is called a nonequilibrium steady-state (NESS).

Biological systems rely on bistability and oscillations to regulate key processes. On one hand, switches control the cellular response to an input signal using complex feedback structures and are key players in many cellular processes such as gene expression, membrane trafficking or vesicular transport. An ideal switch should have an all-or-none response and rely on bistability to be robust against input noise. On the other hand, biochemical oscillators set the timing and control of key processes such as circadian rhythms and the cell cycle. These two regimes are associated with first- and second-order phase transitions, respectively. In this thesis, we will characterize the phase transition occurring in biochemical systems.

Thermodynamics states that complex behaviors in chemical reactions networks such as bistability or oscillations can only be achieved far from equilibrium. In equilibrium, the system will relax towards a unique equilibrium steady-state. The behavior of the system depends on its distance from equilibrium, and thus it must undergo a phase transition at some point. For systems in contact with chemical reservoirs, a control parameter for this phase transition can be a difference of chemical potentials between the reservoirs resulting in thermodynamic forces driving the system in a NESS. For instance, in cells the thermodynamic force depends on the difference in chemical potentials between ATP, ADP and P_i.

Nonequilibrium phase transitions have long been studied and still remain less understood than their equilibrium counterparts. The concept of order parameter can be used in a similar way to equilibrium systems. However, it will not describe the irreversible character of a nonequilibrium phase transition. Recent advances in stochastic thermo-

dynamics have emphasized the importance of entropy production as a quantifier for the distance from equilibrium as it vanishes for equilibrium systems and is positive in a nonequilibrium regime. A natural question that arises is whether the entropy production displays a characteristic behavior at phase transitions.

The study of current fluctuations has recently seen some major developments since the discovery of the "thermodynamic uncertainty relation" in 2015. This relation states that the precision of any current has a minimal thermodynamic cost and holds for a large class of systems described by a Markov process. A current of particular importance is the total entropy production. While entropy current fluctuations have a universal lower bound, they are not bounded from above, especially in the vicinity of a phase transition where we expect the fluctuations of thermodynamic quantities to diverge. The main objective of the work outlined in this thesis is to provide a comprehensive characterization of nonequilibrium phase transitions. In particular, we are interested in the behavior of the entropy production at first- and second-order phase transitions.

Chapter 2: Basics

We review the basic principles of stochastic dynamics, stochastic thermodynamics and chemical reaction networks. Assuming an underlying theory of statistical mechanics, we start with a closed system in contact with a heat bath and identify mesostates. Assuming that the system equilibrates locally within each mesostate, the dynamics become Markovian at the mesoscopic level. A closed system will reach thermal equilibrium, thus, the transition rates between the mesostates must fulfill a detailed balance condition that involves the free energies of the mesostate. By separating the closed system into a core system and chemical reservoirs, we can describe the dynamics of open systems. By tuning the chemical potentials of the reservoirs accordingly, an open system can achieve a NESS contrary to its closed counterpart. We then identify thermodynamic observables such as heat, work or entropy production at the level of individual trajectories in such a way that the expected macroscopic observables are recovered.

We then review the dynamics of chemical reaction networks. With the assumption that chemical reactions are Poisson processes, we define transition rates and introduce the chemical master equation (CME), which describes the dynamics of a chemical reaction network. In general, it is very difficult to solve this master equation exactly. We thus introduce the stochastic simulation algorithm, an exact procedure to generate stochastic trajectories. We then present a series of approximations of the CME. First, we derive the chemical Langevin equation (CLE) from the CME by approximating Poisson random variables with normal random variables. The CLE can also be expressed as Fokker-Planck equation usually obtained by truncating the chemical Kramers-Moyal equation. We derive the van Kampen's linear noise approximation by linearising the CLE. In the limit of infinitely large system sizes, we obtain the deterministic rate equations.

Finally, we discuss whether trimolecular reactions are physically justified. We show that a trimolecular reaction can have the same dynamics at the stochastic level as a set of bimolecular reactions, where an additional short-lived species must be introduced.

Chapter 3: First-order phase transition in biochemical switches

In this chapter, we focus on biochemical switches that undergo a first-order phase transition upon activation. From a deterministic perspective, this phase transition is associated with bistability where two stable steady states can coexist. In contrast, in the stochastic perspective, the steady-state is unique and is associated with a bimodal density distribution. At first-order phase transitions, the entropy production rate is known to be discontinuous. We show that the fluctuations of the entropy production have an exponential volume-dependence in the vicinity of the bistable point. Our results obtained for a paradigmatic model of biochemical switches, Schlögl's model, can be applied to a large class of nonequilibrium systems.

We start with the chemical master equation. Using large deviation theory, we compute the mean rate of entropy production and its associated diffusion coefficient numerically. In the limit of large system sizes, there are two relevant timescales: a fast relaxation to the nearest fixed-point and a slower transition between the fixed-points. Close to the fixed-points, the system can be modeled by stationary Gaussian processes. For the slower dynamics, we coarse-grain the CME into a two-state model, where we show that transition rates are proportional to the exponential of the inverse volume. For this two-state model, we compute an analytical expression for the diffusion coefficient and show that it has an exponential volume-dependence at the bistable point, where the exponential prefactor is given by the height of the effective potential barrier separating the two fixed-points.

Computing diffusion coefficients for a two-state model is straightforward, however, the solution is expressed in terms of the coarse-grained transition rates. The derivation of these rates starting with CME can be challenging, as shown in the appendices of this chapter. As an alternative approach, we derive an analytical expression for the diffusion coefficient starting with the Fokker-Planck equation. For bistable systems, the two-term truncated chemical Kramers-Moyal equation does only correctly describe the dynamics on the shorter timescale and, moreover, does not have the correct stationary distribution. In our case, we consider a Fokker-Planck equation with an effective diffusion term that leads to the same stationary distribution as the CME and correctly describes the dynamics on the larger timescale.

Chapter 4: Second-order phase transition in biochemical oscillators

Biochemical oscillations are ubiquitous in living organisms. In an autonomous system, not influenced by an external signal, they can only occur out of equilibrium. We show that they emerge through a generic nonequilibrium phase transition, with a characteristic qualitative behavior at criticality. We characterize the critical behavior of the entropy production current for three versions of known models: the Brusselator, the activator-inhibitor model, and a model for KaiC oscillations. We find that the first derivative of the mean rate of entropy production has a discontinuity at the critical point, where the fluctuations diverge as a power-law with the volume.

We then discuss metrics for the precision of biochemical oscillations. First, we consider the standard measure of precision, the number of coherent oscillations, which is the number of periods for which different stochastic realizations remain coherent with each other. Second, we consider the Fano factor associated with the thermodynamic flux. In light of the thermodynamic uncertainty relation, the Fano factor is a natural observable to relate precision with thermodynamics. It has been proposed as an observable that can quantify the precision of biochemical oscillators. Since the Fano factor can be small even when there are no biochemical oscillations, we argue that the number of coherent oscillations is more appropriate to quantify the precision of biochemical oscillations.

Concluding perspective

The main goal of this thesis was to describe the behavior of the entropy production current and its fluctuations at first- and second-order phase transitions in homogenous chemical reactions networks. For chemical reaction networks with spatial dynamics, where chemical reactions occur locally and diffusion spreads the chemical species, the characterization of the entropy production remains an open question. Surprisingly, the total entropy production in such systems remains analytical at the phase transition.

We investigated the precision of biochemical oscillations in autonomous systems. As a future direction, it would be interesting to study biochemical oscillators under the influence of an external periodic signal. For these systems, the precision of oscillations can, in principle, be infinite. Moreover, it has been shown that current fluctuations can vanish, thus, violating the original thermodynamic uncertainty relation. Another unexpected way to improve the precision of oscillations is to consider finite reservoirs. We have recently found that they can outperform their ideal counterparts.

1 Introduction

Thermodynamics is a universal theory for describing exchange processes and started with the development of heat engines that can deliver mechanical work using heat reservoirs. In 1824, Carnot proposed a theoretical engine consisting of a set of ideal operations, the Carnot cycle. This engine transfers heat from a warm region to a cold region, converting some of this heat into mechanical work [1]. Thanks to experiments by Joule, the equivalence of heat and mechanical energy was established, and around 1850 Clausius and Thomson stated the two laws of thermodynamics. The first law is the principle of the conservation of energy, it relates the change in internal energy of a system with the extracted work and heat exchanged with the environment. The second law as formulated by Clausius states that "heat can never pass from a colder to a warmer body without some other change, connected therewith, occurring at the same time". In an irreversible cycle, the heat transferred is always greater than the extracted work. This uncompensated heat is now called the total change in entropy production.

Standard thermodynamics is limited to macroscopic systems in equilibrium. Onsager established a first formulation of nonequilibrium thermodynamics in the linear regime, i.e for small perturbations close to equilibrium. He derived universal symmetries, his reciprocal relations, which couples currents to thermodynamic forces with phenomenological coefficients [2, 3]. The connection between thermodynamics and chemical reaction network was pioneered by Prigogine who proposed the concept of local equilibrium that enables the description of irreversible processes using equilibrium quantities [4]. In the last twenty years, nonequilibrium thermodynamics has substantially evolved with stochastic thermodynamics [5–7]. This framework extends the concepts of standard thermodynamics such as work, heat or entropy to the level of individual trajectories. It can describe a large class of small fluctuating systems coupled to a heat bath and arbitrarily far from equilibrium. One of the major successes of this theory is the description of biological processes. For instance, it has been successfully applied to cytoskeletal motors such as kinesin or myosin [8–11] or the F_1-ATPase rotary motor [12, 13]. In the latter case, experimental techniques have enabled the observations of individual trajectories at the single-molecule level [14–16]. From such trajectories, efficiencies for the F_1-ATPase could be computed [17–19].

Stochastic thermodynamics has recently seen major developments in understanding the cost of precision in nonequilibrium steady states. At its origin the discovery of the thermodynamic uncertainty relation (TUR) by Barato and Seifert [20] in 2015, it

was later proved by Gingrich and Horowitz [21] in 2017. This relation is an inevitable trade-off between the cost and precision of any current in a thermodynamically consistent system. In other words, small fluctuations can only be achieved by dissipating a minimal amount of entropy production. From a practical perspective, the TUR can be used to infer hidden properties of complex biochemical networks by analyzing experimental data from single-molecule experiments, in particular, the total entropy production. Most remarkably, the TUR implies an upper bound on the efficiencies of molecular motors [22, 23]. Furthermore, it states that a steady-state heat engine reaching Carnot efficiency at finite power must have diverging power fluctuations [24–27]. From a theoretical standpoint, one possibility to realize such an efficient engine would be to operate it in the vicinity of a phase transition, where fluctuations of observables are known to diverge.

In the work outlined in this thesis, we will provide a comprehensive characterization of nonequilibrium phase transitions. We are particularly interested in the behavior of the entropy production current and its fluctuations at first- and second-order phase transitions.

Phase transitions are arguably one of the most interesting phenomena in the field of physics, they have been long studied in the field of thermodynamics since its origin and remain an active field of research in modern statistical physics. Equilibrium phase transitions have a rich history and are well understood from a thermodynamical standpoint [28, 29]. The first classification scheme for the transition between different phases of matter was introduced by Paul Ehrenfest [30] in 1933 after Keesom and Clusiys [31] discovered a phase transition in liquid helium in 1932, the lambda transition. Ehrenfest proposed that the order of a phase transition to be given by the lowest derivative of the free energy, which is discontinuous at the transition. For instance, the liquid-gas transition would classify as a first-order phase transition as it has a discontinuous change in density, which is the inverse of the first derivative of the free energy with respect to pressure. However, Onsager's solution of the two-dimensional Ising model [32] had a free energy derivative with a power-law divergence rather than a discontinuity close to the critical point. This was experimentally confirmed in experiments by Atkins and Edwards who showed that the lambda transition was of a similar type [33]. This phase transition, which inspired Ehrenfest, turned out to be outside of his classification scheme. By the seventies, the classification scheme was simplified to binary classification with first-order and continuous transitions [34].

Nonequilibrium phase transitions have been long studied since the seventies [35–40] but still remain less understood than their equilibrium counterparts from a thermodynamical perspective. They have been originally classified using nonlinear dynamics and take place at bifurcations [41]. These bifurcations occur in biological systems that can achieve rich dynamics such as biochemical switching and oscillations, which are associated with saddle-node and Hopf bifurcation, respectively. These complex behaviors can be observed, for instance, in interlinked GTPases [42–45] or in the MinDE system [46–50], these generic structures can be tuned be as switches or oscillators depending on

their tasks. Their complex behavior can be understood with simple chemical networks introduced by Schlögl in the seventies to study nonequilibrium first- and second-order phase transitions [51]. Subsequently, order parameters with their associated variances have been derived for these systems [35, 39, 52–55].

Concepts from equilibrium systems such as the order parameter can be used to understand nonequilibrium phase transitions. However, these quantities cannot describe their irreversible character. For instance, the entropy production rate is a quantifier for the distance from equilibrium. It vanishes in equilibrium and is positive in a nonequilibrium regime [56]. A natural question that arises is whether the entropy production rate displays a characteristic behavior at phase transitions. This will be the main focus of this work.

In this thesis, we study first- and second-order phase transitions in chemical reaction networks. We especially focus on the behavior of the entropy production current and its fluctuations. In Chapter 2, we review the basic principles of stochastic dynamics, stochastic thermodynamics, and chemical reaction networks. In Chapter 3, we investigate a paradigmatic model of biochemical switches, Schlögl's model, that undergoes a first-order phase transition upon activation. We start with the chemical master equation and show numerically that the diffusion coefficient diverges at the bistable point. By coarse-graining the chemical master equation, we obtain a two-state model, where we show that the transition rates are proportional to the inverse of the system size. We obtain an analytical expression for the diffusion coefficient and show that it has an exponential volume dependence at the critical point, where the exponential prefactor is the height of the effective potential barrier separating the two fixed-points. As an alternative approach, we also compute the diffusion coefficient starting with the Fokker-Planck equation. In Chapter 4, we study a generic second-order phase transition occurring in autonomous biochemical oscillators. We characterize the critical behavior of the entropy production current in three known models: the Brusselator, the activator-inhibitor model and a model for KaiC oscillations. We show that the first derivative of the mean rate of entropy production is discontinuous at the critical point and that fluctuations diverge as a power-law with the volume. We discuss metrics for the precision of biochemical oscillations by comparing two observables, the Fano factor associated with the thermodynamic flux and the number of coherent oscillations. Since the Fano factor can be small in the absence of oscillations, we argue that the number of coherent oscillations is a better quantifier. In addition, we identify a region where the number of coherent oscillations and the current fluctuations both increase. Finally, we conclude in Chapter 5.

2 Basics

2.1 Stochastic thermodynamics for discrete state spaces

In this section, we review the concepts of stochastic thermodynamics in discrete state space following [7, 57].

2.1.1 Closed systems in contact with a heat bath

We start with a closed system weakly coupled to heat bath at inverse temperature β. The system is fully described by its microstate ξ that contains the positions and momenta of all N particles in the system. Each microstate has an energy $\mathcal{H}(\xi)$, which may not be known explicitly. In equilibrium, the probability to observe a microstate is given by

$$p^{\text{can}}(\xi) = \frac{\exp\left[-\beta\mathcal{H}(\xi)\right]}{\mathcal{Z}(\beta)}, \tag{2.1}$$

where the normalization constant is called the canonical partition function

$$\mathcal{Z}(\beta) \equiv \int \frac{\mathrm{d}\xi}{h^{3N}} \exp\left[-\beta\mathcal{H}(\xi)\right]. \tag{2.2}$$

Here, h is an an arbitrary quantity with the dimension of an action. The free energy, internal energy and entropy of the system are given by

$$F = -(1/\beta)\ln\mathcal{Z}(\beta), \quad U = \frac{\partial}{\partial\beta}(\beta F), \quad S(\beta) = \beta^2\frac{\partial}{\partial\beta}F = \beta(U-F), \tag{2.3}$$

respectively.

Solving the full equation of motion at the microscopic scale is neither possible nor of interest. We group the microstates ξ into a mesostate n and describe the dynamics at the mesoscopic level. Furthermore, we assume that each microstate ξ belongs to exactly one mesostate n. In equilibrium, the probability to observe a mesostate n is

$$P^{\text{eq}}(n) = \sum_{\xi\in n} \exp\left[-\beta\left(\mathcal{H}(\xi)-F\right)\right] \equiv \exp\left[-\beta\left(F_n - F\right)\right], \tag{2.4}$$

where we define F_n, the free energy of state n, in the last equality. The internal energy is defined as

$$U_n = \sum_{\xi \in n} P(\xi|n) = \frac{\partial}{\partial \beta} (\beta F_n), \tag{2.5}$$

where

$$P(\xi|n) = \frac{\exp\left[\beta\left(\mathcal{H}(\xi) - F\right)\right]}{P^{\mathrm{eq}}(n)} = \exp\left[-\beta\left(\mathcal{H}(\xi) - F_n\right)\right] \tag{2.6}$$

is the conditional probability to observe a microstate ξ given the mesostate n. The intrinsic entropy is

$$S_n = \sum_{\xi \in n} P(\xi|n) \ln P(\xi|n) = \beta^2 \frac{\partial}{\partial \beta} F_n = \beta(U_n - F_n), \tag{2.7}$$

where the last term is the Shannon entropy of the conditional probability $P(\xi|n)$.

In equilibrium, the likelihood of observing a forward and backward trajectory is the same by definition, i.e. the probability $P(m, \Delta t|n, 0)$ of observing the system in state n at time $t = 0$ and then later in state m at time t is the same as $P(t, \Delta t|m, 0)$. Therefore, it follows that

$$P(m, t|n, 0)P^{\mathrm{eq}}(n) = P(n, t|m, 0)P^{\mathrm{eq}}(m). \tag{2.8}$$

For a small interval Δt and using Eq. (2.4), we obtain the local detailed balance relation

$$\frac{P^{\mathrm{eq}}(m)}{P^{\mathrm{eq}}(n)} = \frac{w_{nm}}{w_{mn}} = \exp\left[-\beta\left(F_m - F_n\right)\right] = \exp\left[-\beta\left(E_m - E_n\right) + \left(S_m - S_n\right)\right], \tag{2.9}$$

where we define the transition rate from state n to m as

$$w_{nm} \equiv \lim_{\Delta t \to 0} \frac{P(m, t|n, 0)}{\Delta t}. \tag{2.10}$$

We want to describe the dynamics of $n(t)$ as a stochastic process. For that purpose, we will further assume that the dynamics of the microstates within each mesostate is much faster than the dynamics between the mesostates. In other words, we assume that the distribution of microstate within a mesostate relaxes into a local equilibrium distribution on a timescale much faster than the timescale for transitions between the mesostates. As a consequence, the dynamics of the mesostate $n(t)$ becomes memoryless, i.e. the probability for a transition from n to m does not depend on the time spent in state n and states visited before n. Under this assumption of timescale separation, the dynamics of the mesostates can be described by a Markov process with the transition rates defined in Eq. (2.10). The time evolution of the probability $P(n, t)$ to find the system in state n is governed by the master equation, which reads

$$\frac{\partial}{\partial t} P(n, t) = \sum_m \left[P(m, t)w_{mn} - P(n, t)w_{nm}\right]. \tag{2.11}$$

Assuming the the system is ergodic, the probability $P(n,t)$ will converge in the long time limit to the equilibrium distribution $P^{\mathrm{eq}}(n)$, Eq. (2.4), which is time invariant. Indeed, plugging Eq. (2.4) into Eq. (2.11) and using Eq. (2.9) leads to

$$0 = \sum_m \left[P^{\mathrm{eq}}(m)w_{mn} - P^{\mathrm{eq}}(n)w_{nm}\right]. \tag{2.12}$$

In equilibrium, every terms in the sums vanish, i.e. there is no net probability flow across any link nm.

2.1.2 Open systems in contact with a heat bath and chemical reservoirs

Without external control, a closed system in contact with a heat bath will always relax towards equilibrium. In contrast, open systems can achieve a nonequilibrium steady-state (NESS) by extracting energy from their environment. For example, biological systems rely on energy carriers such as adenosine triphosphates (ATP) and extract their energy by hydrolyzing these molecules. Here, we consider an open system coupled to chemical reservoirs. We group the main system and the chemical reservoirs into one large closed supersystem following the ideas from [58–60]. Chemical reservoirs are fully characterized by their chemical potential μ_α and we assume that they are much larger than the main system, i.e. we assume timescale separation between the dynamics of the reservoirs and the dynamics of the main system. From this assumption, the main system can be driven into a NESS by tuning the chemical potentials μ_α accordingly.

The local detailed balance for the supersystem, Eq. (2.9), can be written as

$$\frac{w_{nm}}{w_{mn}} = \exp\left[-\beta\left(F_m - F_n - \sum_\alpha d^\alpha_{nm}\mu_\alpha\right)\right]. \tag{2.13}$$

Here, we split the total free energy of the system into the free energy difference of the main system, $(F_m - F_n)$, and the change in free energy of the reservoirs $\sum_\alpha d^\alpha_{nm}\mu_\alpha$. In a transition from n to m, molecules of species α at a chemical potential μ_α are supplied by the reservoirs ($d_{nm} < 0$) or released in the reservoirs ($d_{nm} > 0$).

Analogously to Eq. (2.12), we can write a master equation for the probability distribution $P(n,t)$ of an open system as

$$\frac{\partial}{\partial t}P(n,t) = \sum_m \left[P(m,t)w_{mn} - P(n,t)w_{nm}\right] = \sum_m j_{mn}(t), \tag{2.14}$$

where we denote

$$j_{nm}(t) \equiv P(m,t)w_{mn} - P(n,t)w_{nm} \tag{2.15}$$

the net probability current from n to m. The Perron-Frobenius theorem [61] states that in the long time limit, $P(n,t)$ will tend towards a unique stationary distribution $P^{\rm ss}(n)$, which is obtained by solving the linear system of equations

$$0 = \sum_m [P^{\rm ss}(m)w_{mn} - P^{\rm ss}(n)w_{nm}] \equiv \sum_m j^{\rm ss}_{mn}, \tag{2.16}$$

where $j^{\rm ss}_{mn}$ is the stationary probability current. In general, these currents do not vanish in a NESS.

2.1.3 Thermodynamics along an individual trajectory

Stochastic thermodynamics is a framework that extends the concepts of classical thermodynamics such as work, heat or entropy to the level of individual trajectories under nonequilibrium conditions [5–7].

In a closed system, Sekimoto [62] proposed that a change in internal energy along an individual trajectory must be compensated by a corresponding change in the energy of the heat bath. Specifically, the first law associated with a transition from n to m reads

$$\Delta E_{nm} = -Q_{nm} \tag{2.17}$$

where $Q_{nm} > 0$ corresponds to heat dissipated in the heat bath. The total entropy change has three contributions. First, there is an increase in entropy of the bath

$$\Delta S^{\rm bath}_{nm} = \beta Q_{nm} \tag{2.18}$$

due to the heat Q_{nm} dissipated. Second, there is a change of intrinsic entropy ΔS_{nm}, defined in Eq. (2.7), due to a change of mesostate. The last contribution is the change in stochastic entropy [63]

$$\Delta S^{\rm stoch}_{nm}(t) \equiv \ln \frac{P(n,t)}{P(m,t)}. \tag{2.19}$$

If the probability distribution $P(n,t)$ is not stationary, the stochastic entropy $S^{\rm stoch}(t) \equiv -\ln P(n,t)$ will change even when the system remains in the same mesostate. These last two contributions can be combined as the change in system entropy

$$\Delta S^{\rm sys}_{nm}(t) \equiv \Delta S_{nm} + \Delta S^{\rm stoch}_{nm}(t). \tag{2.20}$$

Consequently, the total entropy change for a transition from n to m can be written as as

$$\Delta S^{\rm tot}_{nm}(t) = \beta Q_{nm} + \Delta S^{\rm sys}_{nm}(t) = \ln \frac{P(n,t)w_{nm}}{P(m,t)w_{mn}}, \tag{2.21}$$

where the last equality follows from Eqs. (2.9) and (2.17).

For an open system in contact with a heat bath and chemical reservoirs, the first law (2.17) along a transition from n to m becomes

$$E_{nm} = -W^{\text{chem}}_{mn} - Q_{nm}, \tag{2.22}$$

where we identify the chemical work performed on the reservoirs as

$$W^{\text{chem}}_{mn} \equiv \sum_{\alpha} d^{\alpha}_{nm}\mu_{\alpha}. \tag{2.23}$$

In a NESS, the total entropy production change along a stochastic trajectory $n(t)$ can be written using Eq. (2.21) as

$$\Delta S^{\text{tot}}\left[n(t)\right] = \sum_{n,m} Z_{nm}(t) \ln \frac{P(n,t)w_{nm}}{P(m,t)w_{mn}}, \tag{2.24}$$

where $Z_{nm}(t)$ is a random variable that counts the number of transition from n to m in the time interval $[0,t]$. The mean rate of entropy production is [64]

$$\sigma = \sum_{mn} P(n,t)w_{nm} \ln \frac{P(n,t)w_{nm}}{P(m,t)w_{mn}} = \sum_{n<m} j_{nm} \ln \frac{P(n,t)w_{nm}}{P(m,t)w_{mn}} \geq 0. \tag{2.25}$$

In a stationary state, the contributions from the intrinsic and stochastic entropy are not extensive in time. Consequently, the rate of entropy production in the steady state is the same as the rate of entropy production in the heat bath, leading to

$$\sigma = \sum_{n<m} j^{\text{ss}}_{nm} \ln \frac{P(n)w_{nm}}{P(m)w_{mn}} = \sum_{n<m} j^{\text{ss}}_{nm} \ln \frac{w_{nm}}{w_{mn}}, \tag{2.26}$$

where we obtain the last equality using Eq. (2.16).

2.2 Chemical reaction networks

We consider a closed system with volume Ω consisting of a mixture of molecular species $\{X_1, ..., X_N\}$. We assume that the chemical mixture is in a dilute solution, which means that the molecular species are entirely surrounded by the solvent. The solvent has several roles. First, it mediates the heat and volume exchanges. Second, it screens the reactive species X_i from each other, thus, ensuring that any configuration of the system is locally stable. Finally, we assume that the system is in contact with a heat bath at inverse temperature β and to be well-stirred, in other words, the system does not have any spatial inhomogeneities in the number of species.

The dynamical state of the system is given by a vector $\mathbf{X}(t) = \left(N_{X_1}(t), ..., N_{X_N}(t)\right)$, where $N_{X_i}(t)$ is the number of X_i molecules at time t. The state of the system evolves

when a chemical reaction occurs. The set of chemical reactions $R_1, ..., R_M$ forms a chemical reaction network. A typical reaction in channel R_ρ can be written as

$$\sum_{i=1}^{N} s_i X_i \xrightarrow{k_\rho} \sum_{i=1}^{N} r_i X_i \tag{2.27}$$

Here, k_ρ is the reaction rate, s_i and r_i are stoichiometric coefficients that characterize, respectively, the numbers of molecules of the associated species being consumed and produced by the reaction. A reaction of type (2.27) changes the state of the system $\mathbf{n}$ into $\mathbf{n} + \boldsymbol{\nu}$, where we define the stoichiometric vector $\boldsymbol{\nu}$ as $\nu_i \equiv r_i - s_i$. Thermodynamic consistency requires all reactions to be reversible. For readability, we will assume that each reaction R_ρ has a corresponding reverse reaction $R_{\rho'}$, which is included in the set of chemical reactions $R_1, ..., R_M$. This reverse reaction changes the state of the system $\mathbf{n}$ into $\mathbf{n} - \boldsymbol{\nu}$ at rate $k_{\rho'}$. In addition, thermodynamic consistency imposes a constraint on the ratio of rates $k_\rho / k_{\rho'}$, which must fulfill the local detailed balance relation (2.9).

2.2.1 Rates of chemical reactions

We want to describe the dynamics of the chemical networks as a stochastic process. For that purpose, we must compute the rate at which chemical reactions occur. We define the propensity function $a_\rho(\mathbf{X})$ as follows. For an infinitesimal time increment $\mathrm{d}t$,

$$\begin{aligned} a_\rho(\mathbf{X})\mathrm{d}t \equiv{}& \text{the probability, for a given state } \mathbf{X}(t), \text{ that one reaction } R_\rho \\ & \text{will take place inside } \Omega \text{ in the time interval } [t, t + \mathrm{d}t],\ \rho = 1, .., M.. \end{aligned} \tag{2.28}$$

In a chemical reaction of type (2.27), the reacting molecules must first bind together to form a complex and then dissociate into product molecules. We assume that the timescale for the formation and dissociation of this complex is much faster than the reaction and that this process is energetically neutral. These assumptions ensure that such a chemical reaction can be described by a one-step process in a thermodynamically consistent way [65]. The probability for a collision involving s_i molecules X_i should be proportional to the number of possibilities to choose s_i molecules out of the total numbers N_{X_i} for species i, i.e.

$$\propto N_{X_i} \left(N_{X_i} - 1\right) \left(N_{X_i} - 2\right) ... \left(N_{X_i} - s_i + 1\right) = \frac{N_{X_i}!}{\left(N_{X_i} - s_i\right)!}. \tag{2.29}$$

Once the complex is formed, the chemical reaction takes place. Here we will consider elementary reactions where this rate is given by a constant k_ρ given by the Eyring-Kramers equation [66–68]. By summing over all reactants, we obtain the propensity function

$$a_\rho(\mathbf{X}) = \Omega k_\rho \prod_{i=1}^{N} \frac{N_{X_i}!}{\Omega^{s_i} \left(N_{X_i} - s_i\right)!}. \tag{2.30}$$

This expression is also known as the mass-action rate for elementary reactions. One can easily check that the propensity function $a_\rho(\mathbf{X})$ has the expected units, namely, a number of collisions per unit time. For a chemical reaction network, we add the contributions from all reaction channels. The resulting dynamics are then described by the chemical master equation (CME) defined as follows.

Chemical master equation For a set of M chemical reactions $R_1, ..., R_M$, the time evolution of the probability $P(\mathbf{X}, t)$ to find the system in state $\mathbf{X}$ is given by

$$\frac{\partial}{\partial t}P(\mathbf{X},t) = \sum_{\rho=1}^{M}\Big[a_\rho(\mathbf{X}-\boldsymbol{\nu})P(\mathbf{X}-\boldsymbol{\nu},t) - a_\rho(\mathbf{X})P(\mathbf{X},t)\Big]. \tag{2.31}$$

Chemical Kramers-Moyal equation For sufficiently large systems such that the variables $\mathbf{X}$ can be considered as real numbers (i.e. $\mathbf{X} \gg 1$). If we assume that the function $f_\rho(\mathbf{X}) = a_\rho(\mathbf{X})p(\mathbf{X}, t)$ is analytical, we can expand it as a Taylor series,

$$f_\rho(\mathbf{X}-\boldsymbol{\nu}) = f_\rho(\mathbf{X}) + \sum_{n=1}^{\infty}\sum_{\substack{m_1,\dots,m_N \\ m_1+\dots+m_N=n}}^{n} \frac{1}{m_1!\dots m_N!}\left(-\nu_{\rho 1}\right)^{m_1}\dots\left(-\nu_{\rho N}\right)^{m_N}\frac{\partial^n f_\rho(\mathbf{X})}{\partial N_{X_1}^{m_1}\dots\partial N_{X_N}^{m_N}}. \tag{2.32}$$

We substitute this expression into the CME, Eq. (2.31), and obtain the chemical Kramers-Moyal equation

$$\frac{\partial}{\partial t}P(\mathbf{X},t) = \sum_{n=1}^{\infty}\sum_{\substack{m_1,\dots,m_N \\ m_1+\dots+m_N=n}}^{n} \frac{(-1)^n}{m_1!\dots m_N!}\frac{\partial^n}{\partial N_{X_1}^{m_1}\dots\partial N_{X_N}^{m_N}}\sum_{\rho}^{M}\left(\nu_{\rho 1}^{m_1}\dots\nu_{\rho N}^{m_N}\right)a_\rho(\mathbf{X})P(\mathbf{X},t). \tag{2.33}$$

2.2.2 Simulation of the chemical master equation

In 1976 Gillespie [69, 70] proposed an exact procedure to generate realizations of $\mathbf{X}(t)$ called the stochastic simulation algorithm (SSA). Given a state $\mathbf{X}(t)$ at time t, the SSA selects at what time $t + \tau$ the next reaction takes place and which reaction R_i it is. Since τ and i are random variables, we want to compute the joint probability density function, $p(\tau, i|\mathbf{X}, t)$. Specifically, $p(\tau, i|\mathbf{X}, t)\mathrm{d}\tau$ is the probability, for a given $\mathbf{X}$, that the next reaction takes place in the time interval $[t + \tau, t + \tau + \mathrm{d}\tau]$ and is reaction R_i. As the reactions are independent Poisson processes, this joint probability is given by

$$p(\tau, i|\mathbf{X}, t) = \mathrm{e}^{-a_{\mathrm{tot}}(\mathbf{X})\tau}a_i(\mathbf{X}), \tag{2.34}$$

where

$$a_{\text{tot}}(\mathbf{X}) \equiv \sum_{j=1}^{N} a_j(\mathbf{X}). \tag{2.35}$$

In practice, the SSA consists of the following computational steps:

1. Compute the propensities $a_j(\mathbf{X})$ at time t and draw two random numbers, u_1 and u_2, from a uniform distribution over the interval $[0, 1]$. The next reaction will occur at time [71]

   $$\tau = \frac{1}{a_{\text{tot}}} \ln\left(\frac{1}{1-u_1}\right) \tag{2.36}$$

   and it will be reaction R_i, where i is chosen such that the relation

   $$\sum_{j=1}^{i-1} a_j(\mathbf{X}) < u_2 a_{\text{tot}}(\mathbf{X}) \leq \sum_{j=1}^{i} a_j(\mathbf{X}) \tag{2.37}$$

   is satisfied.

2. Update the state of the system $\mathbf{X} \rightarrow \mathbf{X} + \nu$ and the time $t \rightarrow t + \tau$.

3. Return to step 1 or end the simulation.

2.2.3 Approximations of the chemical master equation

In general, it is very difficult to solve the chemical master equation exactly and simulations of large systems can be computationally inefficient. For that purpose, we present a series of approximations of the CME.

The chemical Langevin equation

For a sufficiently large numbers of molecules X_i in a volume Ω, Gillespie [72] proposed in 2001 an approximation called tau-leaping, where τ is a leap time chosen such that the propensity functions fulfills two conditions. First, we require τ to be small enough such that the change in the state of the system $\mathbf{X}$ can be neglected, i.e. the propensity functions satisfy

$$a_\rho(\mathbf{X}) \approx \text{constant in } [t, t+\tau], \quad \forall \rho \qquad \text{(first leap condition)}. \tag{2.38}$$

Since the propensity functions $a_\rho(\mathbf{X})$ are constant in the time interval $[t, t+\tau]$, we can assume that all reactions occurring during this time interval are independent of each other. Therefore, the number of R_ρ reactions occurring in $[t, t+\tau]$ is given by a set of

Poison random variables $\mathcal{P}(a_\rho(\mathbf{X}(\mathbf{t})))$. Each of these reactions will increase the number of molecules $\mathbf{X}$ by $\boldsymbol{\nu}_\rho$. Thus, the concentrations of X_i molecules at time $t + \tau$ is given by

$$\mathbf{X}(t+\tau) = \mathbf{X}(t) + \sum_\rho \boldsymbol{\nu}_\rho \mathcal{P}\left(a_\rho(\mathbf{X}(t))\right). \tag{2.39}$$

Second, we require τ to be large enough such that the number of reactions in each channel R_ρ is large, i.e.

$$\left\langle \mathcal{P}\left(a_\rho(\mathbf{X}(t))\right)\right\rangle = a_\rho(\mathbf{X})(t)\tau \gg 1, \quad \forall\rho \qquad \text{(second leap condition)}. \tag{2.40}$$

If Eq. (2.40) is satisfied, we can apply the Central Limit Theorem and approximate the Poisson random variables $\mathcal{P}\left(a_\rho(\mathbf{X}(t))\right)$ by normal random variables $\mathcal{N}_\rho(\mu, \sigma^2)$. Using the property $\mathcal{N}(\mu, \sigma^2) = \mu + \sigma^2\mathcal{N}(0, 1)$, we can write Eq. (2.39) as

$$\mathbf{X}(t+\tau) = \mathbf{X}(t) + \sum_\rho \boldsymbol{\nu}_\rho a_\rho(\mathbf{X}(t)) + \boldsymbol{\nu}_\rho\sqrt{a_\rho(\mathbf{X}(t))}\mathcal{N}_\rho(0,1)\sqrt{\tau}, \tag{2.41}$$

From now on, we consider τ as an infinitesimal time that satisfies both leap conditions and formally replace it with $\mathrm{d}t$. We also write $\mathcal{N}_\rho(0, 1)$ as $\mathcal{N}_\rho(t)$, where $\mathcal{N}_\rho(t)$ and $\mathcal{N}_{\rho'}(t')$ are statistically independent if either $\rho \neq \rho'$ or $t \neq t'$. From Eq. (2.41), we obtain

$$\mathbf{X}(t+\mathrm{d}t) = \mathbf{X}(t) + \sum_\rho \boldsymbol{\nu}_\rho a_\rho(\mathbf{X}(t)) + \boldsymbol{\nu}_\rho\sqrt{a_\rho(\mathbf{X}(t))}\mathcal{N}_\rho(t)\sqrt{\mathrm{d}t}, \tag{2.42}$$

which has the canonical form of a Langevin equation and is called the Chemical Langevin equation (CLE). Eq. (2.42) can be transformed in an equivalent "white-noise" form

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}t} = \sum_\rho \boldsymbol{\nu}_\rho a_\rho(\mathbf{X}(t)) + \boldsymbol{\nu}_\rho\sqrt{a_\rho(\mathbf{X}(t))}, \boldsymbol{\xi}(t) \tag{2.43}$$

where $\boldsymbol{\xi}(t)$ is a vector of uncorrelated, statistically independent Gaussian white noises, i.e.

$$\left\langle \xi_i(t)\right\rangle = 0, \quad \text{and} \quad \left\langle \xi_i(t)\xi_j(t')\right\rangle = \delta_{ij}\delta(t-t'). \tag{2.44}$$

There are two caveats in the derivation of the CLE. First, the system must admit an infinitesimal time $\mathrm{d}t$ that satisfies both leap conditions Eqs. (2.38) and (2.40), which is not a trivial statement. In fact, one can easily find systems for which that requirement cannot be made, and thus not be described the CLE. However, Gillespie [73] proved in 2009 that both leap conditions can always be satisfied as long as the number of molecules is sufficiently large, specifically in the thermodynamic limit where the number of molecules $\mathbf{X}$ and the system size Ω both go to infinity while the concentrations $\mathbf{X}\Omega^{-1}$ remain finite. Second, the CLE cannot accurately describe "rare" events as we have approximated a

Poisson process with a Gaussian process. Sample values close to the mean value will be accurate, however, the tails of these distributions will differ by many order of magnitudes. Consequently, the CLE will underestimate the likelihood of rare events and only describe the "typical" behavior of a chemical system.

For sufficiently large systems, we can assume that the variables $\mathbf{X}$ are real numbers. From Eq. (2.43), the probability density function for $\mathbf{X}(t)$ obeys a forward Fokker-Planck [74]

$$\frac{\partial}{\partial t}P(\mathbf{X},t) = \sum_{i=1}^{N}\frac{\partial}{\partial N_{x_i}}\left[-\sum_{\rho}\nu_{\rho i}a_\rho(\mathbf{X}) + \frac{1}{2}\sum_{j=1}^{N}\frac{\partial}{\partial N_{x_j}}\nu_{\rho i}\nu_{\rho j}a_\rho(\mathbf{X})\right]P(\mathbf{x},t) \tag{2.45}$$

Eq. (2.45) can also be obtained by truncating the Kramers-Moyal chemical master Eq. (2.33) at $n = 2$. However, the physical justification behind this approximation is not straightforward. Kurtz showed in a lengthy derivation that the difference in concentrations between the solution of the chemical master Eq. (2.31) and that the chemical Langevin Eq. (2.43) is proportional to $\log\Omega/\Omega$, thus, vanishes in the thermodynamic limit [75–77]. In a different contex, Van Kampen [78, 79] proposed in 1961 an alternative way of deriving Eq. (2.45), the "system-size expansion". Starting with Kramers-Moyal chemical master Eq. (2.33) and a change of variable from $\mathbf{X}(t)$ to $\mathbf{x}(t) = \mathbf{X}(t)/\Omega$, Van Kampen performs a series expansion around the noiseless solution of Eq. (2.43) and obtain Eq. (2.45) in the lowest order in $\Omega^{-1/2}$.

Deterministic rate equations

We consider the thermodynamic limit, where the number of molecules $\mathbf{X}$ and the system size Ω both go to infinity in way such that the concentrations $\mathbf{x} = \mathbf{X}/\Omega$ remain finite. The propensity functions, Eq. (2.27), diverges linearly with the system size. In the thermodynamic limit, we define the deterministic transition rates

$$\widetilde{a}_\rho(\mathbf{x}) \equiv \Omega^{-1}a_\rho(\mathbf{X}) \tag{2.46}$$

By considering only the higher order terms in Ω, we can approximate Eq. (2.46) as

$$\widetilde{a}_\rho(\mathbf{x}) \approx k_\rho\prod_j x_j^{s_j}. \tag{2.47}$$

The CLE, Eq. (2.42), with concentrations reads

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \sum_{\rho}\boldsymbol{\nu}_\rho\left(\mathbf{X}(t)\right)\widetilde{a}_\rho(\mathbf{x})(t) + \Omega^{-1/2}\sum_{\rho}\boldsymbol{\nu}_\rho\sqrt{a_\rho(\mathbf{x}(t))}\boldsymbol{\xi}(t) \tag{2.48}$$

In the deterministic limit ($\Omega \to \infty$), the second term can be neglected and we obtain the rate equations

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \sum_{\rho}\boldsymbol{\nu}_\rho(\mathbf{x})\widetilde{a}_\rho(\mathbf{x})(t). \tag{2.49}$$

Linear noise approximation

The linear noise approximation (LNA) is usually obtained by linearizing the two-term truncated chemical Kramers-Moyal equation, Eq. (2.33). Here, we will follow a different approach and compute the LNA as an approximation of the CLE following Wallace *et al.* [80]. The chemical Langevin Eq. (2.48) and the rate Eq. (2.49) differ by a term proportional to $\Omega^{-1/2}$. We make the ansatz that the solution $\mathbf{x}(t)$ of the CLE will differ from the solution $\hat{\mathbf{x}}(t)$ of the rate equations by a term proportional to $\Omega^{-1/2}$, i.e.

$$\mathbf{x}(t) = \hat{\mathbf{x}}(t) + \Omega^{-1/2}\boldsymbol{\xi}(t). \tag{2.50}$$

By substituting Eq. (2.50) into Eq. (2.42) and replacing $\hat{\mathbf{x}}(t + \mathrm{d}t) - \hat{\mathbf{x}}(t)$ with $\sum_\rho \widetilde{a}_\rho(\hat{\mathbf{x}}(t))$, we obtain an equation for the stochastic function $\boldsymbol{\xi}(t)$ that reads

$$\begin{aligned}\boldsymbol{\xi}(t + \mathrm{d}t) - \boldsymbol{\xi}(t) = \Omega^{1/2} \sum_\rho \boldsymbol{\nu}_\rho \left[\widetilde{a}_\rho\left(\hat{\mathbf{x}} + \Omega^{-1/2}\boldsymbol{\xi}(t)\right) - \widetilde{a}_\rho(\hat{\mathbf{x}}(t))\right]\mathrm{d}t \\ + \sum_\rho \boldsymbol{\nu}_\rho \sqrt{\widetilde{a}_\rho\left(\hat{\mathbf{x}} + \Omega^{-1/2}\boldsymbol{\xi}(t)\right)}\mathcal{N}_\rho(t)\sqrt{\mathrm{d}t}.\end{aligned} \tag{2.51}$$

We expand the propensity function around $\hat{\mathbf{x}}(t)$,

$$\widetilde{a}_\rho\left(\hat{\mathbf{x}} + \Omega^{-1/2}\boldsymbol{\xi}(t)\right) = \widetilde{a}_\rho\left(\hat{\mathbf{x}}\right) + \Omega^{-1/2} \sum_{i=1}^{N} y_{\rho i}(t)\xi_k(t), \tag{2.52}$$

where we define

$$y_{\rho i}(t) = \left.\frac{\partial \widetilde{a}_\rho(\mathbf{x})}{\partial x_i}\right|_{\mathbf{x}=\hat{\mathbf{x}}(t)}. \tag{2.53}$$

Finally, by substituting Eq. (2.52) into Eq. (2.51) and retaining only the higher order terms, we obtain

$$\boldsymbol{\xi}(t + \mathrm{d}t) - \boldsymbol{\xi}(t) = \sum_\rho \sum_{i=1}^{N} \boldsymbol{\nu}_\rho y_{\rho i}(t)\xi_k(t)\mathrm{d}t + \boldsymbol{\nu}_\rho \sqrt{\widetilde{a}_\rho\left(\hat{\mathbf{x}}\right)}\mathcal{N}_\rho(t)\sqrt{\mathrm{d}t}, \tag{2.54}$$

van Kampen's LNA [78, 79]. The main difference to the chemical Langevin Eq. (2.43) is that the stochastic term in Eq. (2.54) is independent of $\xi(t)$.

2.2.4 Physical justification for the propensities functions of multimolecular reactions

The expressions for the propensities can be computed for any reaction described by Eq. (2.27), where the number of reactants can be in principle an arbitrary number.

However, we assumed that reactions are instantaneous and occur in a dilute system. Thus, the likelihood of a reaction taking place with three or more reactants is very low [71]. In fact, chemical reactions are mainly of two types: unimolecular, where one molecule is transformed into products; or bimolecular, where two molecules collide and are transformed into products. First, we will justify the form of the propensity function, Eq. (2.30), for unimolecular or bimolecular reactions. Then, we will show under which conditions a trimolecular reaction can be approximated as a series of bimolecular reactions.

Unimolecular reactions A reaction of the form $X_1 \longrightarrow \ldots$ is described by a Poisson process. The probability that a random X_1 molecule decays inside a volume Ω in a time interval $\mathrm{d}\tau$ is equal to some constant k_ρ times $\mathrm{d}\tau$. By summing over the total number of molecules N_{X_1}, the propensity function is $a_\rho^{\mathrm{uni}}(N_{X_1}) = k_\rho N_{X_1}$.

Bimolecular reactions For reactions of the form $X_1 + X_2 \longrightarrow \ldots$, computing the reaction rate can be more challenging. Gillespie originally proposed a simple kinetic argument [69, 70], which reads as follows. We assume that X_1 and X_2 are hard spheres of radii r_1 and r_2, respectively. A collision occurs whenever the distance between a X_1 molecule and a X_2 molecule is equal to $r_1 + r_2$. We pick an arbitrary pair of $X_1 - X_2$ molecules and denote v_{12} the velocity of X_1 relative to X_2. In the time interval $\mathrm{d}\tau$, the molecule X_1 will cover a reaction volume $\mathrm{d}\Omega_{\mathrm{coll}} = \pi r_{12}^2 v_{12} \mathrm{d}\tau$ relative to molecule X_2. A reaction will take place if the molecule X_2 is inside this volume $\mathrm{d}\Omega_{\mathrm{coll}}$. Assuming that the system is well-mixed, the probability to find a molecule X_2 in the collision volume is simply given by the ratio $\mathrm{d}\Omega_{\mathrm{coll}}/\Omega$. Averaging over the velocity distributions of the X_1 and X_2 molecules, this ratio is given on average by

$$\overline{\mathrm{d}\Omega_{\mathrm{coll}}/\Omega} = \pi r_{12}^2 \overline{v}_{12} \Omega^{-1} \mathrm{d}\tau, \tag{2.55}$$

where $v = \overline{v}_{12}$ is the average relative speed. We denote q_ρ, the probability that an $X_1 - X_2$ collision will result in a chemical reaction. By summing over the total number of molecules $N_{X_1} + N_{X_2}$, we obtain the propensity function

$$a_\rho^{\mathrm{bis}}(N_{X_1}, N_{X_2}) = q_\rho \left(\pi r_{12}^2 \overline{v}_{12} \Omega^{-1} \right) N_{X_1} N_{X_2}. \tag{2.56}$$

Note that this result is only valid when the solvent is a dilute gas.

For a generalization of Eq. (2.56) to the case where only the reactants are dilute, we refer the reader to Sections 2.7 and 3.8 of [81]. Under the assumptions that X_1 and X_2 are well-mixed and dilute into a bath of much smaller solvent molecules, Gillespie and Seitaridou show explicitly that

$$a_\rho^{\mathrm{bis}}(N_{X_1}, N_{X_2}) = k_\rho \left(\frac{4\pi r_{12}^2 D_{12} \overline{v}_{12} \Omega^{-1}}{4D_{12} + r_{12} \overline{v}_{12}} \right) N_{X_1} N_{X_2}, \tag{2.57}$$

where D_{12} is the sum of the diffusion coefficients of X_1 and X_2 molecules. For fast diffusion ($4D_{12} \gg r_{12}\overline{v}_{12}k_\rho$), Eq. (2.57) reduces to Eq. (2.56). We have thus shown that the propensity function, defined in Eq. (2.30) for bimolecular reactions is justified.

Trimolecular reactions In Chapters 3 and 4, we will consider a trimolecular reaction of the form

$$2X_1 + X_2 \xrightarrow{\mathrm{k}_0} \dots \tag{2.58}$$

Using Eq. (2.30), the propensity functions can be written as

$$a_0^{\mathrm{tri}}(N_{X_1}, N_{X_2}, N_{X_3}) = k_0 N_{X_1} N_{X_2} N_{X_3}/\Omega^2. \tag{2.59}$$

and the CME reads

$$\frac{\partial}{\partial t}P(\mathbf{X}, t) = \frac{k_0}{\Omega^2}\left(\mathcal{E}_{X_1}^{+}\mathcal{E}_{X_1}^{+}\mathcal{E}_{X_2}^{+} - 1\right) N_{X_1}\left(N_{X_1} - 1\right) N_{X_2} P(\mathbf{X}, t) \tag{2.60}$$

where $\mathbf{X} = \left(N_{X_1}, N_{X_2}\right)$. For readability, we introduce the step operators

$$\begin{aligned} \mathcal{E}_{X_1}^{\pm} P(\mathbf{X}, t) &\equiv P(N_{X_1} \pm 1, N_{X_2}, t), \quad & \mathcal{E}_{X_1}^{\pm} N_{X_1} &\equiv \left(N_{X_1} \pm 1\right), \\ \mathcal{E}_{X_2}^{\pm} P(\mathbf{X}, t) &\equiv P(N_{X_1}, N_{X_2} \pm 1, t), \quad & \mathcal{E}_{X_2}^{\pm} N_{X_2} &\equiv \left(N_{X_2} \pm 1\right), \end{aligned} \tag{2.61}$$

However, the volume scaling of Eq. (2.59) is in contradiction with the assumption that the reactants are dilute. The likelihood of the trimolecular reaction (2.58) taking place is very unlikely and would thus make the propensities function, Eq. (2.59), nonphysical. In addition, the probability that three randomly chosen hard-spheres come into contact in the time interval $\mathrm{d}\tau$ will be proportional to $(\mathrm{d}\tau)^2$, such a reaction cannot be described by a Poisson process.

It turns out that a series of bimolecular reactions can result in an effective trimolecular reaction. We introduce an additional species, Y_1, and decompose Eq. (2.58) into a set of bimolecular reactions

$$\begin{aligned} X_1 + X_1 &\underset{k_{-1}/\epsilon}{\overset{k_1}{\rightleftharpoons}} Y_1, \\ X_2 + Y_1 &\xrightarrow{k_2/\epsilon} \dots. \end{aligned} \tag{2.62}$$

We will prove that the dynamics of this network is the same as the trimolecular reaction (2.58) at the stochastic level using a perturbation theory with $\epsilon \ll 1$, i.e. the reactions with rates k_{-1}/ϵ and k_2 are much faster than the one with rate k_1. The first line of Eq. (2.62) implies a conservation law:

$$\overline{N_{X_1}} \equiv N_{X_1} + 2N_{Y_1}, \tag{2.63}$$

i.e. the variable $\overline{N_{X_1}}$ does not change when the reactions with rate k_1 and k_{-1}/ϵ take place. For reasons which will become apparent later, we replace N_{X_1} with $\overline{N_{X_1}}$ in the

CME, the state of the system is now given by $\overline{\mathbf{X}} = \left(\overline{N_{X_1}}, N_{X_2}\right)$ and N_{Y_1}. We write the CME on the slow timescale $\tau \equiv t\epsilon$,

$$\frac{\partial}{\partial \tau} P_\epsilon(\overline{\mathbf{X}}, N_{Y_1}, t) = \left(\frac{1}{\epsilon^2}\mathcal{L}^{\text{fast}} + \frac{1}{\epsilon}\mathcal{L}^{\text{slow}}\right) P_\epsilon(\overline{\mathbf{X}}, N_{Y_1}, \tau), \tag{2.64}$$

Here, we split the dynamics of the CME into two parts. A fast operator

$$\mathcal{L}^{\text{fast}} \equiv k_{-1}\left(\mathcal{E}_{Y_1}^{+} - 1\right) N_{Y_1} + k_2\left(\mathcal{E}_{\overline{X_1}}^{+}\mathcal{E}_{X_1}^{+}\mathcal{E}_{X_2}^{+}\mathcal{E}_{Y_1}^{+} - 1\right) N_{X_2}N_{Y_1}, \tag{2.65}$$

where the step operator of $\mathcal{E}_{Y_1}^{\pm}$ is defined analogously to Eq. (2.61) and a slow operator

$$\mathcal{L}^{\text{slow}} \equiv k_1\left(\mathcal{E}_{Y_1}^{-} - 1\right)\left(\overline{N_{X_1}} - 2N_{Y_1}\right)\left(\overline{N_{X_1}} - 2N_{Y_1} - 1\right). \tag{2.66}$$

We write the solution of Eq. (2.64) as the perturbation series

$$P_\epsilon(\overline{\mathbf{X}}, N_{Y_1}, \tau) = P_0(\overline{\mathbf{X}}, N_{Y_1}, t) + \epsilon P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) + \epsilon^2 P_2(\overline{\mathbf{X}}, N_{Y_1}, \tau). \tag{2.67}$$

Substituting this expression into Eq. (2.64) and equating terms of equal powers in ϵ, we obtain three equations:

$$\begin{aligned}
O\left(1/\epsilon^2\right) &: \mathcal{L}^{\text{fast}} P_0(\overline{\mathbf{X}}, N_{Y_1}, \tau) = 0, \\
O\left(1/\epsilon\right) &: \mathcal{L}^{\text{fast}} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) = -\mathcal{L}^{\text{slow}} P_0(\overline{\mathbf{X}}, N_{Y_1}, \tau), \\
O\left(1\right) &: \mathcal{L}^{\text{fast}} P_2(\overline{\mathbf{X}}, N_{Y_1}, \tau) = \frac{\partial}{\partial \tau} P_0(\overline{\mathbf{X}}, N_{Y_1}, \tau) - \mathcal{L}^{\text{slow}} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau).
\end{aligned} \tag{2.68}$$

Order $1/\epsilon^2$ In the limit $\epsilon \to 0$, we expect that Y_1 will be a short-lived species. Thus, the solution at order $1/\epsilon^2$ will tend to

$$P_0(\overline{\mathbf{X}}, N_{Y_1}, \tau) = P_0(\overline{\mathbf{X}}, \tau) P_0(N_{Y_1}) = P_0(\overline{\mathbf{X}}, \tau)\delta_{N_{Y_1},0}\,, \tag{2.69}$$

where

$$\sum_{\overline{\mathbf{X}}} P_0(\overline{\mathbf{X}}, \tau) = 1 \tag{2.70}$$

and $\delta_{N_{Y_1},0}$ is the Kronecker-delta function. Here, we note that $\mathcal{L}^{\text{fast}} P_0(\overline{\mathbf{X}}, N_{Y_1}, \tau) = 0$ reduces to $\mathcal{L}^{\text{fast}} P_0(N_{Y_1}) = 0$.

Order $1/\epsilon$ By inserting Eq. (2.69) into the second line of Eq. (2.68), we obtain by summing over N_{Y_1}

$$\sum_{N_{Y_1}} \mathcal{L}^{\text{fast}} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) = P_0(\overline{\mathbf{X}}, \tau) k_1 \sum_{N_{Y_1}} \left(\overline{N_{X_1}} - 2N_{Y_1}\right)\left(\overline{N_{X_1}} - 2N_{Y_1} - 1\right)\left(\mathcal{E}_{Y_1}^- - 1\right)\delta_{N_{Y_1},0}. \tag{2.71}$$

In the limit $\epsilon \to 0$, the species N_{N_Y} is short-lived, thus, the first-order solution of Eq. (2.67) can be split as follows,

$$P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) = P_0(\overline{\mathbf{X}}, \tau)\delta_{N_{Y_1},0}. \tag{2.72}$$

By substituting this expression into the left hand side of Eq. (2.71), we obtain by summing over N_{Y_1}

$$\begin{aligned}\sum_{N_{Y_1}} \mathcal{L}^{\text{fast}} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) &= \sum_{N_{Y_1}} \Big[k_{-1}\left(\mathcal{E}_{Y_1}^+ - 1\right) + k_2\left(\mathcal{E}_{\overline{X_1}}^+ \mathcal{E}_{\overline{X_1}}^+ \mathcal{E}_{X_2}^+ \mathcal{E}_{Y_1}^+ - 1\right) N_{X_2} \\ &\qquad \Big] N_{Y_1} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) \\ &= \left[k_{-1} + k_2\left(\mathcal{E}_{\overline{X_1}}^+ \mathcal{E}_{\overline{X_1}}^+ \mathcal{E}_{X_2}^+ - 1\right) N_{X_2}\right] P_0(\overline{\mathbf{X}}, \tau) \sum_{N_{Y_1}} N_{Y_1}\delta_{N_{Y_1},0}.\end{aligned} \tag{2.73}$$

Assuming that $P_0(\mathbf{X}, \tau) > 0$ and using the fact that $(\mathcal{E}_{Y_1}^-) = -\left(\mathcal{E}_{Y_1}^+ - 1\right)\mathcal{E}_{Y_1}^-$, Eq. (2.71) becomes

$$\begin{aligned}\sum_{N_{Y_1}} \Big\{ &\left(\mathcal{E}_{Y_1}^+ - 1\right)\left[k_{-1}N_{Y_1} - k_1\mathcal{E}_{Y_1}^-\left(\overline{N_{X_1}} - 2N_{Y_1}\right)\left(\overline{N_{X_1}} - 2N_{Y_1} - 1\right)\right] \\ &+ k_2\left(\mathcal{E}_{\overline{X_1}}^+ \mathcal{E}_{\overline{X_1}}^+ \mathcal{E}_{X_2}^+ - 1\right) N_{X_2} N_{Y_1}\Big\}\delta_{N_{Y_1},0} = 0.\end{aligned} \tag{2.74}$$

The expressions proportional to k_2 vanishes due to $\delta_{N_{Y_1},0}$. Finally, we obtain the relation

$$\sum_{N_{Y_1}} N_{Y_1}\delta_{N_{Y_1},0} = \frac{k_1}{k_{-1}}\overline{N_{X_1}}\left(\overline{N_{X_1}} - 1\right), \tag{2.75}$$

Order 1 To solve the third line of Eq. (2.68), we note that the left hand side vanishes,

$$\sum_{N_{Y_1}} \mathcal{L}^{\text{fast}} P_2(\overline{\mathbf{X}}, N_{Y_1}, \tau) = 0, \tag{2.76}$$

i.e. the fast operator $\mathcal{L}^{\text{fast}}$ only has a non-zero contribution for the lower order in ϵ in the perturbation serie, Eq. (2.67), the higher order terms evolve on the slow timescale τ.

The right hand side is

$$\sum_{N_{Y_1}} \left[\frac{\partial}{\partial \tau} P_0(\overline{\mathbf{X}}, N_{Y_1}, \tau) - \mathcal{L}^{\text{slow}} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) \right] = 0. \tag{2.77}$$

The first term reduces to $\partial/\partial\tau P_0(\overline{\mathbf{X}}, \tau)$. Using Eq. (2.72), we can write the second term in the sum as

$$\begin{aligned} \sum_{N_{Y_1}} \mathcal{L}^{\text{slow}} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) &= k_1 \sum_{N_{Y_1}} \left(\mathcal{E}_{Y_1}^- - 1 \right) \left(\overline{N_{X_1}} - 2N_{Y_1} \right) \left(\overline{N_{X_1}} - 2N_{Y_1} - 1 \right) P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) \\ &= k_1 k_2 \left(\mathcal{E}_{X_1}^+ \mathcal{E}_{X_1}^+ \mathcal{E}_{X_2}^+ - 1 \right) N_{X_2} P_0(\overline{\mathbf{X}}, \tau) \sum_{N_{Y_1}} N_{Y_1} \delta_{N_{Y_1},0}. \end{aligned} \tag{2.78}$$

Using Eq. (2.75), we obtain

$$\sum_{N_{Y_1}} \mathcal{L}^{\text{slow}} P_1(\overline{\mathbf{X}}, N_{Y_1}, \tau) = k_1 k_2 \left(\mathcal{E}_{X_1}^+ \mathcal{E}_{X_1}^+ \mathcal{E}_{X_2}^+ - 1 \right) N_{X_2} \overline{N_{X_1}} \left(\overline{N_{X_1}} - 1 \right) P_0(\overline{\mathbf{X}}, \tau). \tag{2.79}$$

Combining these expressions, we obtain an effective CME

$$\frac{\partial}{\partial \tau} P_0(\overline{\mathbf{X}}, \tau) = \frac{k_1 k_2}{k_{-1}} \left(\mathcal{E}_{X_1}^+ \mathcal{E}_{X_1}^+ \mathcal{E}_{X_2}^+ - 1 \right) N_{X_2} \overline{N_{X_1}} \left(\overline{N_{X_1}} - 1 \right) P_0(\overline{\mathbf{X}}, \tau). \tag{2.80}$$

This equation is equivalent the chemical master Eq. (2.60) if we tune the rates such that $k_0 = k_1 k_2 / k_{-1}$ and in the limit $\epsilon \to 0$.

We have thus shown that a trimolecular reaction of type (2.58) can be decomposed into two biomolecular reactions, Eq. (2.62), and have the same dynamics at the stochastic level assuming timescale separation. Note that our choice of biomolecular reactions is not unique. For a multimolecular reaction with arbitrary molecularity, Wilhelm proposed a systematic decomposition into a set of biomolecular reactions [82].

2.2.5 Open chemical reaction networks

Until now, we have dealt with closed chemical reaction networks where the total number of molecules $\sum_i N_{X_i}$ is a conserved quantity. A closed system will always tend towards a unique equilibrium point in the long time limit. Open chemical reaction networks coupled to chemical reservoirs can achieve a nonequilibrium steady-state.

We consider an open system with volume Ω consisting of a mixture of molecular intermediate species $\{X_1, ..., X_N\}$ and chemostatted species $\{B_1, ..., B_C\}$ supplied at constant concentrations by chemical reservoirs, also called chemostats. We consider that the reservoirs are ideal ones and provide a constant chemical potential μ_j for all chemostatted

species B_j. We can connect the chemical potentials μ_j with the concentrations b_j using the equation for an ideal dilute solution,

$$\mu_j = \mu_0 + \beta^{-1} \ln \frac{b_j}{b_0}, \tag{2.81}$$

where μ_0 is the chemical potential for standard conditions at concentration b_0. A typical reaction can be written as

$$\sum_{i=1}^{N} s_i X_i + \sum_{j=1}^{C} s_j B_j \xrightarrow{k_\rho} \sum_{i=1}^{N} r_i X_i + \sum_{j=1}^{C} r_j B_j \tag{2.82}$$

The dynamical state of the system is given by the vector of intermediate species $\mathbf{X}(t) = \big(N_{X_1}(t), ..., N_{X_N}(t)\big)$, where $N_{X_i}(t)$ is the number of X_i molecules in the system at time t. Note that the state of the system does not depend explicitly on the concentrations of the chemostatted species. The time evolution of the system is described by the CME, Eq. (2.31). The propensity function associated with reaction (2.82) is given by

$$a_\rho(\mathbf{X}) = \Omega k_\rho \prod_{i}^{N} \frac{N_{X_i}!}{\Omega^{s_i} \left(N_{X_i} - s_i\right)!} \prod_{j=1}^{C} b_j^{s_j}. \tag{2.83}$$

For thermodynamic consistency, the ratio of the reaction rate k_ρ and its inverse counterpart must fulfill the local detailed balance relation for open systems, Eq. (2.13).

3 First-order phase transitions in biochemical switches

3.1 Introduction

In this chapter, we focus on biochemical switches that undergo a first-order phase transition upon activation. From a deterministic perspective, this phase transition is associated with bistability where two stable steady states can coexist. In contrast, in the stochastic perspective, the steady-state is unique and is associated with a bimodal density distribution [83, 84]. In the thermodynamic limit, the stochastic system will relax to the more stable fixed-point except at the bistable point [85, 86]. There are two timescales relevant for a biochemical switch: a fast relaxation to the nearest fixed-point and a slower transition between the states, where the coarse-grained transition rates are proportional to the exponential of the inverse volume [87, 88]. In chemical reaction networks, bistability can occur only far from equilibrium since an equilibrium distribution will always be a Poisson distribution distribution, and, thus, have a single peak [77, 89].

The behavior of the entropy production at first- and second-order phase transitions has been investigated in many systems such as chemical networks or nonequilibrium Ising models [84–86, 90–100]. At first-order phase transitions, the entropy production rate has a discontinuity with respect to the thermodynamic force whereas at second-order phase transitions its first derivative has a discontinuity. Recently, it has been shown that the critical fluctuations of the entropy production diverge with a power-law with the volume at a second-order phase transition [98]. For first-order phase transitions, the behavior of entropy production fluctuations has not been investigated yet to the best of our knowledge.

We will show that the fluctuations of the entropy production have an exponential volume-dependence at first-order phase transitions in chemical reaction networks. Our results are obtained for Schlögl's model. First, we compute the entropy fluctuations numerically from the chemical master equation using standard large deviation techniques [101, 102]. Second, we compute the current fluctuations for a coarse-grained two-state model and show that the diffusion coefficient diverges at the bistable point with an exponential prefactor given by the height of the effective potential barrier separating the two fixed-points. As an alternative approach, we derive an analytical expression for the diffusion coefficient starting with the Fokker-Planck equation.

The chapter is organized as follows. In Section 3.2, we introduce Schlögl's model

and define the entropy production. In Section 3.3, we consider the chemical master equation and compute the diffusion coefficient numerically. In Section 3.4, we introduce an effective two-state model and compute an analytical expression for the diffusion coefficient. In Section 3.5, we compute the diffusion coefficient starting with the Fokker-Planck equation. We conclude in Section 3.6.

3.2 Schlögl's model and entropy production

3.2.1 Model definition

The Schlögl model is a paradigmatic model for biochemical switches [51, 84]. It consists of a chemical species X in a volume Ω. The external bath contains two chemical species A and B at fixed concentrations a and b, respectively. The set of chemical reactions is

$$\begin{aligned} 2X + A &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 3X, \\ B &\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} X, \end{aligned} \tag{3.1}$$

where k_1, k_{-1}, k_2 and k_{-2} are transition rates. The system is driven out of equilibrium due to a difference of chemical potential between A and B, which is written as $\Delta\mu \equiv \mu_A - \mu_B$. A cycle in which an X molecule is created with rate k_1, and then degraded with rate k_{-2} leads to the consumption of a substrate A and generation of a product B. The thermodynamic force associated with this cycle is

$$\Delta\mu \equiv \ln \frac{k_{-2}k_1 a}{k_{-1}k_2 b}, \tag{3.2}$$

where the temperature T and Boltzmann's constant k_B are set to 1 throughout this chapter.

3.2.2 Entropy production

Along a stochastic trajectory $n(t)$, where n labels the state with n molecules of species X, the entropy production change of the heat bath can be identified as [6]

$$\Delta s^{\text{bath}} = Z_B(t) \ln \frac{k_{-2}}{k_2 b} + Z_A(t) \ln \frac{k_1 a}{k_{-1}}. \tag{3.3}$$

Here, $Z_A(t)$ and $Z_B(t)$ are random variables which count transitions in the A and B channel, respectively. For example, $Z_B(t)$ increases by one if a B is produced, which

happens if a reaction with rate k_{-2} takes place. Likewise, it decreases by one if a B is consumed, which happens if a reaction with rate k_2 takes place, i.e.

$$\begin{aligned} 2X + A &\xrightarrow{k_1} 3X & Z_A &\to Z_A + 1\,, \\ 3X &\xrightarrow{k_{-1}} 2X + A & Z_A &\to Z_A - 1\,, \\ X &\xrightarrow{k_{-2}} B & Z_B &\to Z_B + 1\,, \\ B &\xrightarrow{k_2} X & Z_B &\to Z_B - 1\,. \end{aligned} \tag{3.4}$$

In this chapter, we focus on Δs^{bath}, the extensive part of the total entropy production $\Delta s^{\mathrm{tot}} \equiv \Delta s^{\mathrm{bath}} + \Delta s$. The remaining part is the change in stochastic entropy [63]

$$\Delta s = -\ln p_{n_t}(t) + \ln p_{n_0}(0), \tag{3.5}$$

where $p_{n_t}(t)$ is the probability to find the system in state n_t at time t.

Using Eqs. (3.3) and (3.4), we can write the mean entropy production rate per volume in the steady-state as [6]

$$\begin{aligned} \sigma &\equiv \lim_{t\to\infty} \frac{\langle \Delta s^{\mathrm{bath}} \rangle}{t\Omega} \\ &= \lim_{t\to\infty} \left[\frac{\langle Z_B(t) \rangle}{t\Omega} \ln \frac{k_{-2}}{k_2 b} + \frac{\langle Z_A(t) \rangle}{t\Omega} \ln \frac{k_1 a}{k_{-1}} \right] \\ &= J_B \Delta\mu. \end{aligned} \tag{3.6}$$

In the steady-state, the mean flux density of B molecules

$$J_B \equiv \lim_{t\to\infty} \frac{\langle Z_B(t) \rangle}{t\Omega} = \lim_{t\to\infty} \frac{\langle Z_A(t) \rangle}{t\Omega} \tag{3.7}$$

is equal to the flux of A molecules consumed.

We can quantify the fluctuations of B molecules with the diffusion coefficient

$$\begin{aligned} D_B &\equiv \lim_{t\to\infty} \frac{\langle Z_B(t)^2 \rangle - \langle Z_B(t) \rangle^2}{2t\Omega} \\ &= D_A \equiv \lim_{t\to\infty} \frac{\langle Z_A(t)^2 \rangle - \langle Z_A(t) \rangle^2}{2t\Omega}, \end{aligned} \tag{3.8}$$

where we prove the second equality, $D_B = D_A$, in Appendix 3.A. Specifically, we show there that $Z_A(t)$ and $Z_B(t)$ have the same cumulants.

Finally, using Eqs. (3.3) and (3.8) we obtain the diffusion coefficient associated with the entropy production in the heat bath as

$$\begin{aligned} D_\sigma &\equiv \lim_{t\to\infty} \frac{\langle (\Delta s^{\mathrm{bath}})^2 \rangle - \langle \Delta s^{\mathrm{bath}} \rangle^2}{2t\Omega} \\ &= D_B \left(\ln \frac{k_{-2}}{k_2 b} \right)^2 + D_A \left(\ln \frac{k_1 a}{k_{-1}} \right)^2 \\ &+ \lim_{t\to\infty} \left[\frac{\langle Z_A(t) Z_B(t) \rangle - \langle Z_A(t) \rangle \langle Z_B(t) \rangle}{2t\Omega} \right] 2 \left(\ln \frac{k_{-2}}{k_2 b} \right) \left(\ln \frac{k_1 a}{k_{-1}} \right) \\ &= D_B \Delta\mu^2. \end{aligned} \tag{3.9}$$

Since the stochastic entropy production is not extensive in time, this diffusion coefficient is equal to the one for the total entropy production.

3.3 Chemical master equation

3.3.1 Stationary solution

The state of the system is fully determined by the total number n of X molecules. The time evolution of $P(n,t)$, which is the probability to find the system in state n at time t, is governed by the chemical master equation (CME)

$$\begin{aligned} \frac{\partial}{\partial t} P(n,t) &= f_{n-1} P(n-1,t) + g_{n+1} P(n+1,t) - (f_n + g_n) P(n,t) \\ &\equiv \mathcal{L}_0 P(n,t). \end{aligned} \tag{3.10}$$

Here, we define the rate parameters as

$$\begin{aligned} f_n &= \alpha_n^+ + \beta_n^+ \equiv \frac{a k_1 n(n-1)}{\Omega} + b k_2 \Omega, \\ g_n &= \alpha_n^- + \beta_n^- \equiv \frac{k_{-1} n(n-1)(n-2)}{\Omega^2} + k_{-2} n. \end{aligned} \tag{3.11}$$

The system can reach a nonequilibrium steady-state with a distribution written as P_n. The analytical solution for this steady-state probability distribution reads

$$\begin{aligned}\frac{P_n}{P_0} &= \prod_{i=0}^{n-1} \frac{f_i}{g_{i+1}}, \\ &= \frac{f_0}{g_1}\sqrt{\frac{f_1 g_1}{f_n g_n}} \\ &\quad \exp\left[\frac{1}{2}\ln\left(\frac{f_1}{g_1}\right) + \sum_{i=2}^{n-1}\ln\left(\frac{f_i}{g_i}\right) + \frac{1}{2}\ln\left(\frac{f_n}{g_n}\right)\right] \\ &\quad (n \geq 3),\end{aligned} \tag{3.12}$$

where the normalization is given by

$$P_0 = 1 - \sum_{j=1}^{\infty} P_j. \tag{3.13}$$

For a large number of states ($\Omega \to \infty$), we can write $x = n/\Omega$ as a continuous variable and approximate the exponential by an integral using the trapezium rule, which is valid if f_n/g_n is bounded. Without loss of generality, we will choose parameters such that the fixed-points are far enough from the boundary, which means that the probability to find the system close to $x = 0$ is negligible. Note that the case where the solution does not vanish at the boundary is discussed in details in [88].

From Eq. (3.12), the continuous steady-state distribution can be written as

$$p(x) \propto \exp\left[-\Omega\left(\phi_0(x) + \frac{1}{\Omega}\phi_1(x)\right)\right] \tag{3.14}$$

where we define the non-equilibrium potential

$$\phi_0(x) \equiv -\int_0^x \mathrm{d}y\ \ln\left(\frac{f(y)}{g(y)}\right), \tag{3.15}$$

and

$$\phi_1(x) \equiv -\frac{1}{2}\ln\frac{1}{f(x)g(x)}. \tag{3.16}$$

Here, we have defined the total transition rates

$$\begin{aligned} f(x) &\equiv \alpha^+(x) + \beta^+(x) = ak_1x^2 + k_2b, \\ g(x) &\equiv \alpha^-(x) + \beta^-(x) = k_{-1}x^3 + k_{-2}x. \end{aligned} \tag{3.17}$$

In the deterministic limit ($\Omega \to \infty$), we obtain the equation of the time evolution of the density,

$$\bar{x} \equiv \sum_n nP(n,t)/\Omega \tag{3.18}$$

as

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = f(\bar{x}) - g(\bar{x}) \tag{3.19}$$

from the chemical master Eq. (3.10). In the steady-state, this equation has three solutions (x_-, x_0, x_+). Bistability occurs when all solutions are real, we order the fixed-points as follows: $0 < x_- < x_0 < x_+$, where $x_\pm$ are stable (i.e. $f'(x_\pm) < g'(x_\pm)$) and x_0 is unstable (i.e. $f'(x_0) > g'(x_0)$).

3.3.2 Behavior of the entropy production at the phase transition

The mean entropy production rate, defined in Eq. (3.6), can be written using Eq. (3.17) as

$$\sigma = \Delta\mu \int_0^\infty \mathrm{d}x \left[\beta^-(x) - \beta^+(x)\right] p(x), \tag{3.20}$$

In the thermodynamic limit, the stochastic system will relax to the more stable fixed-point except at the bistable point [85, 86]. Consequently, the rate of entropy production will be discontinuous with respect to the thermodynamic force at the bistable point. Specifically, with increasing $\Delta\mu$, σ will jump from $[\beta^-(x_-) - \beta^+(x_-)]\,\Delta\mu$ to $[\beta^-(x_+) - \beta^+(x_+)]\,\Delta\mu$, which are the rates of entropy production at the two fixed-points x_- and x_+, respectively.

We now derive an expression for the fluctuations of the entropy production. We follow an approach based on large deviation theory [101–103] and considered for Brownian ratchets in [104]. We want to compute the cumulants related to the number of produced B molecules. They are obtained through the scaled cumulant generating function (SCGF)

$$\alpha(\lambda) \equiv \lim_{t\to\infty} \frac{1}{t} \ln\langle \mathrm{e}^{\lambda Z_B}\rangle \tag{3.21}$$

where Z_B is the time-integrated current of B molecules defined in Eq. (3.4). Note that $\alpha(\lambda)$ is unrelated to the transition rate defined in Eq. (3.11). From now on, we will drop the t dependence on Z_B for readability. Expanding the generating function yields

$$\alpha(\lambda) = \Omega J_B \lambda + \Omega D_B \lambda^2 + O(\lambda^3) \tag{3.22}$$

The SCGF can be obtained by considering the moment generating function

$$\begin{aligned} g(\lambda, t) &\equiv \langle \mathrm{e}^{Z_B}\rangle \\ &= \sum_n g(\lambda, n, t) P(n,t) \end{aligned} \tag{3.23}$$

where we define

$$\begin{aligned} g(\lambda, n, t) &\equiv \langle \mathrm{e}^{\lambda Z_B} | n(t) = n \rangle \\ &= \sum_{Z_B} \mathrm{e}^{\lambda Z_B} P(n, Z_B, t), \end{aligned} \tag{3.24}$$

which is conditioned on the final state of the trajectory $n(t)$. The time evolution of this quantity is given by

$$\frac{\partial}{\partial t} g(\lambda, n, t) = \sum_m \mathcal{L}_{nm}(\lambda) g(\lambda, m, t). \tag{3.25}$$

where $\mathcal{L}(\lambda)$ is the tilted operator. Note that for $\lambda = 0$, it is identical to the operator generating the time evolution of the probability distribution in Eq. (3.10)

We want to specify $\mathcal{L}(\lambda)$ for the chemical master Eq. (3.10). First, we write the time evolution of the probability distribution as

$$P(n, Z_B, t) = \sum_m w_{mn} P(m, Z_B - d_{mn}, t) - w_{nm} P(n, Z_B, t) \tag{3.26}$$

where w_{mn} is the transition rate from state m to state n and d_{mn} is the distance matrix which characterize how Z_B changes during a transition. In the case of the CME, Eq. (3.26) reduces to

$$\begin{aligned} P(n, Z_B, t) = &\alpha^-_{n+1} P(n+1, Z_B, t) \\ &+ \alpha^-_{n-1} P(n-1, Z_B, t) \\ &+ \beta^-_{n+1} P(n+1, Z_B + 1, t) \\ &+ \beta^+_{n-1} P(n-1, Z_B - 1, t) \\ &- \left(\alpha^+_n + \beta^+_n + \alpha^-_n + \beta^-_n\right) P(n, Z_B, t) \end{aligned} \tag{3.27}$$

where the rates $\alpha^\pm_n$ and $\beta^\pm_n$ are given by Eq. (3.11). Using Eq. (3.24), we can write

$$\begin{aligned} \frac{\partial}{\partial t} g(\lambda, n, t) = &\sum_m w_{mn} \sum_{Z_B} \mathrm{e}^{\lambda Z_B} P(m, Z_B - d_{mn}, t) \\ &- w_{nm} g(\lambda, n, t). \end{aligned} \tag{3.28}$$

With a change of variable, we obtain

$$\begin{aligned} \frac{\partial}{\partial t} g(\lambda, n, t) &= \sum_m w_{mn} \sum_{Y_B} \mathrm{e}^{\lambda (Y_B + d_{mn})} P(m, Y_B, t) \\ &\quad - w_{nm} g(\lambda, n, t) \\ &= \sum_m \mathrm{e}^{\lambda d_{mn}} w_{mn} g(\lambda, m, t) - w_{nm} g(\lambda, n, t), \end{aligned} \tag{3.29}$$

which specifies the tilted operator $\mathcal{L}(\lambda)$ defined in Eq. (3.25).

The cumulants can be obtained by solving the eigenvalue equation $\mathcal{L}(\lambda)\mathbf{Q}(\lambda) = \alpha(\lambda)\mathbf{Q}(\lambda)$, where

$$\mathcal{L}(\lambda) = \mathcal{L}_0 + \mathcal{L}_1\lambda + \mathcal{L}_2\lambda^2 + O(\lambda^3) \tag{3.30}$$

and the distribution

$$\mathbf{Q}(\lambda) = \mathbf{P} + \mathbf{Q}_1\lambda + \mathbf{Q}_2\lambda^2 + O(\lambda^3). \tag{3.31}$$

Sorting by orders of λ, we obtain

$$\begin{aligned} \mathcal{L}_0\mathbf{P} &= 0, \\ \mathcal{L}_0\mathbf{Q}_1 + \mathcal{L}_1\mathbf{P} &= \Omega J_B\mathbf{P}, \\ \mathcal{L}_1\mathbf{Q}_1 + \mathcal{L}_0\mathbf{Q}_2 + \mathcal{L}_2\mathbf{P} &= \Omega D_B\mathbf{P} + \Omega J_B\mathbf{Q}_1. \end{aligned} \tag{3.32}$$

We multiply these equations with $\langle 1|$ on the left hand side and note that $\mathbf{P}$ is normalized, i.e. $\langle 1 \,|\, \mathbf{P}\rangle = 1$, where $\langle\cdot|\cdot\rangle$ denotes the standard scalar product. We can compute the mean flux density

$$J_B = \frac{1}{\Omega}\langle 1|\,\mathcal{L}_1\,|\mathbf{P}\rangle = \frac{1}{\Omega}\sum_n \left(\beta_n^- - \beta_n^+\right) P_n \tag{3.33}$$

and the diffusion coefficient

$$\begin{aligned} D_B &= \frac{1}{\Omega}\Big(\langle 1|\,\mathcal{L}_1\,|\mathbf{Q}_1\rangle + \langle 1|\,\mathcal{L}_2\,|\mathbf{P}\rangle\Big) - J_B\,\langle 1\,|\,\mathbf{Q}_1\rangle \\ &= \frac{1}{\Omega}\sum_n \left(\left(\beta_n^- - \beta_n^+\right)(Q_1)_n + \frac{1}{2}\left(\beta_n^- + \beta_n^+\right) P_n\right) \\ &\quad - J_B\sum_n (Q_1)_n. \end{aligned} \tag{3.34}$$

The mean rate of entropy production σ as well as its associated diffusion coefficient D_σ can be evaluated using Eqs. (3.6) and (3.9).

3.3.3 Numerical results

Throughout this paper, we set the parameters to $k_1 = 1, k_2 = 0.2, k_{-2} = 1, a = 1$ and $b = 1$. The transition rate k_{-1} is computed from $\Delta\mu$ and the generalized detailed balance relation Eq. (3.2), where $\Delta\mu$ is a control parameter of the phase transition.

In Fig. 3.1(a), we plot the stationary distribution $p(x)$ which is bimodal in the vicinity of the phase transition ($\Delta\mu^{\mathrm{bi}} \simeq 3.045$). In Fig. 3.1(b), we plot the entropy production rate σ as a function of $\Delta\mu$. With increasing system size, σ gets steeper at the bistable point. In Fig. 3.1(c), we show that the first derivative of σ follows a power-law with an effective prefactor close to 1. In the thermodynamic limit ($\Omega \to \infty$), the entropy production rate becomes discontinuous as shown by Ge and Qian [85, 86].

The diffusion coefficient D_σ reaches a maximum at the bistable point and has an exponential volume-dependence. In Fig. 3.2(a), we compare simulations of the chemical master Eq. (3.10) using Gillespie's algorithm [70] for three increasing sampling times T with the numerical solution obtained by solving the linear system given by Eq. (3.34). The systematic difference between these two methods is due to a limited sampling time. In the next section, we present an effective two-state model and derive an analytical expression for the diffusion coefficient.

3.4 Two-state model

3.4.1 Stationary solution

In the bistable regime, the system has two timescales. First, it will relax towards the nearest stable fixed-points $x_\pm$ and fluctuate around it. Close to the fixed-points, the system can be modeled by stationary Gaussian processes. Specifically, the distribution $p(x)$ can be expanded around its stable fixed-points $x_\pm$ as [84],

$$p(x) \approx \sum_{x_*=(x_-,x_+)} \frac{\mathrm{e}^{-\Omega\phi_0(x_*)}}{\mathcal{Z}_G A_2(x_*)} \exp\left[\frac{-\Omega\phi_0''(x_*)\,(x-x_*)^2}{2}\right], \tag{3.35}$$

where

$$\mathcal{Z}_G \equiv \sum_{x_*=(x_-,x_+)} \frac{\mathrm{e}^{-\Omega\phi_0(x_*)}\sqrt{2\pi}}{A_2(x_*)\sqrt{\left|\phi_0''(x_*)\right|\,\Omega}}. \tag{3.36}$$

Second, as stochastic fluctuations are always present, the system will at some point in time reach the unstable fixed-point x_0 beyond which it can relax towards the other fixed-point. Based on this behavior, the infinite-state system Eq. (3.10) can be coarse-grained into a two-state process between the stable fixed-points $x_\pm$ [105]. The transition rates from $x_\pm$ to $x_\mp$ depend exponentially on the system size and are given explicitly by [87, 88]

$$r_\pm = \frac{\mathrm{e}^{-\Omega[\phi_0(x_0)-\phi_0(x_\pm)]} f(x_\pm)\sqrt{-\phi_0''(x_0)\phi_0''(x_\pm)}}{2\pi\Omega}. \tag{3.37}$$

In Appendix 3.B, we derive these transition rates following Hinch and Chapman [88].

3.4.2 Behavior of the entropy production at the phase transition

We want to characterize the fluctuations of the entropy production for the two-state model. We consider the thermodynamic flux J_B and its associated diffusion coefficient D_B defined

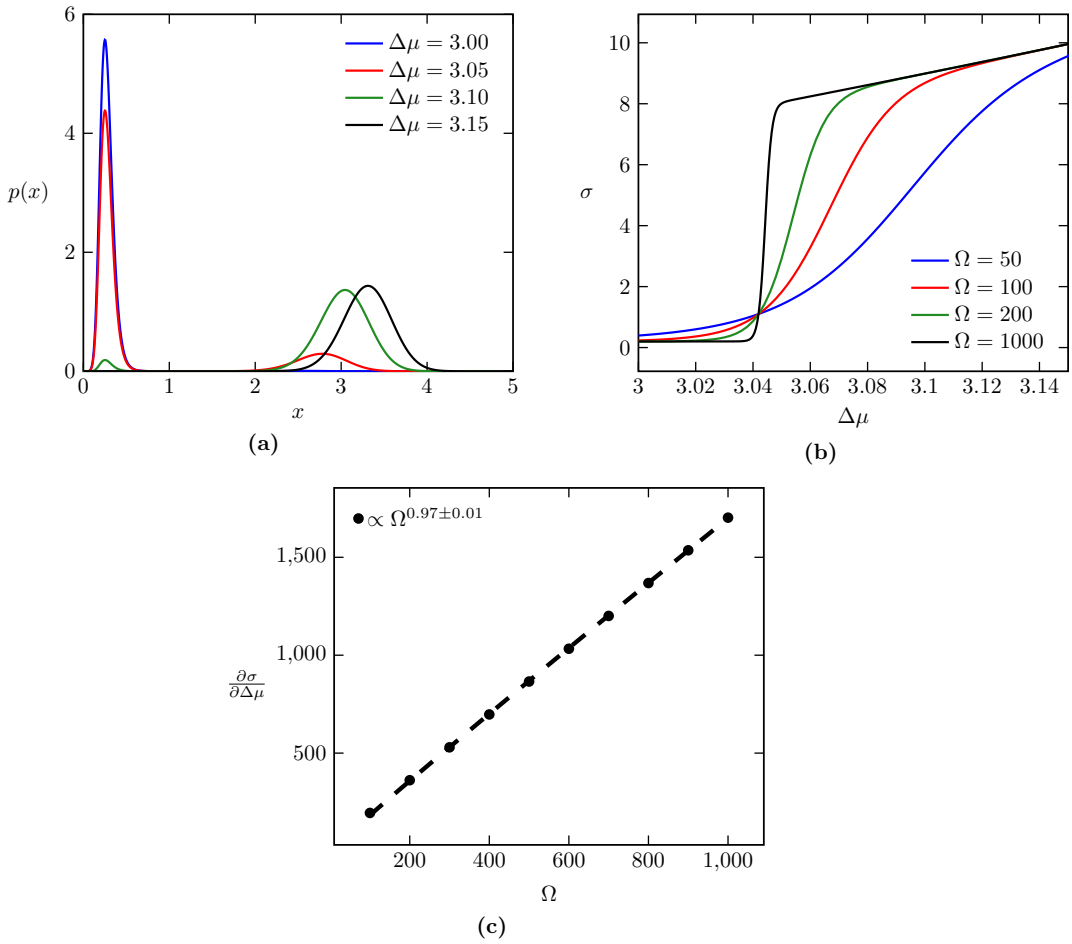

Figure 3.1: Phase transition in the Schlögl model. (a) Stationary distribution of chemical species X for $\Omega = 100$ and different values of $\Delta\mu$. (b) Mean entropy production rate σ as a function of $\Delta\mu$ for different system sizes Ω. (c) Maximum of the first derivative of σ as a function of the system size Ω. Parameters are given in the main text.

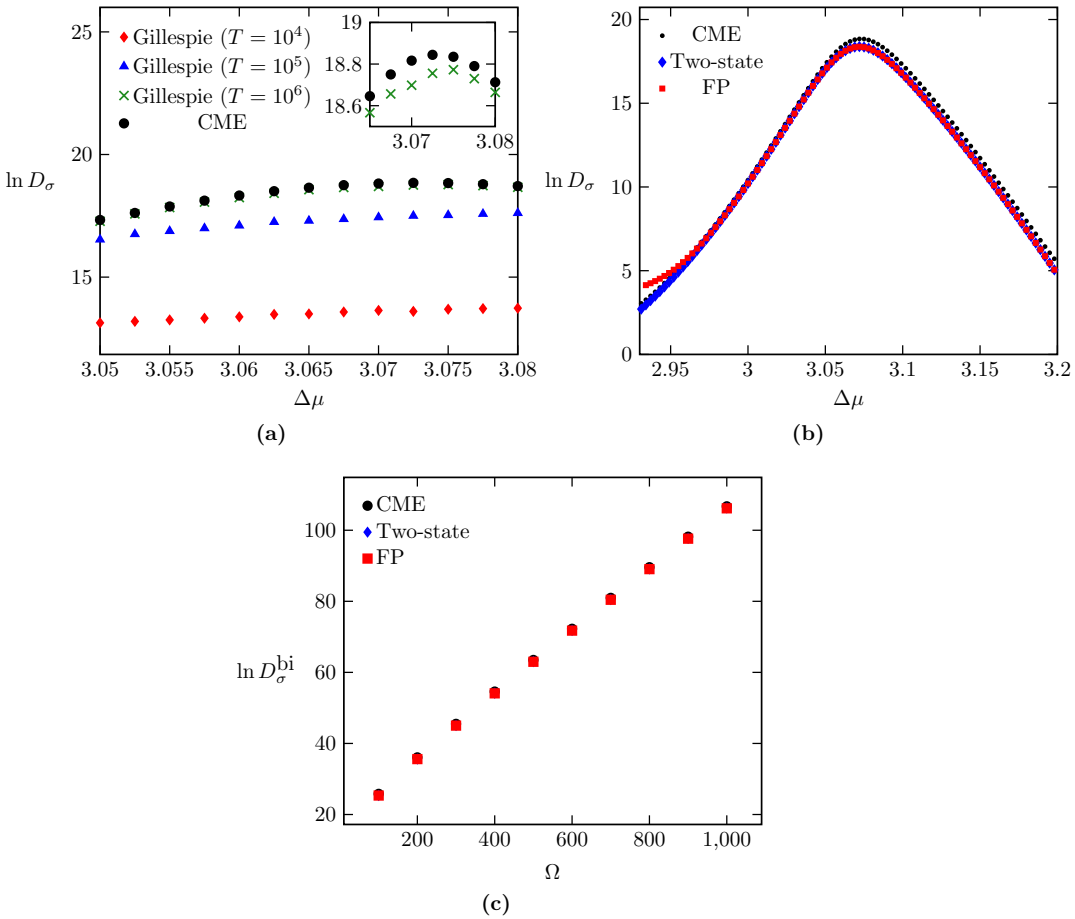

Figure 3.2: Behavior of the diffusion coefficient D_σ close to the bistable point. (a) Diffusion coefficient from simulations using Gillespie's algorithm for 10^4 trajectories with a sampling time of $T = 10^4, 10^5, 10^6$. For the CME, we solve the linear system given by Eq. (3.34) numerically for $\Omega = 100$. (b) Diffusion coefficient for the two-state model obtained by evaluating Eq. (3.43) and for the Fokker-Planck equation (FP) by evaluating Eq. (3.74) (c) Finite-size scaling of the maximum of the diffusion coefficient D_σ^{bi}. The corresponding slopes are given in Table 3.1.

in Eqs. (3.7) and (3.8). There are two contributions to both of these quantities. First, the system can fluctuate around one of the fixed-point, which is modeled by a Gaussian process Eq. (3.35). Within the state $x_\pm$, the thermodynamic flux $J_\pm$ and its diffusion coefficient $D_\pm$ can be computed exactly [102]. Second, the system can jump from state $x_\pm$ to $x_\mp$ on the largest timescale. For simplicity, we will assume that a transition from $x_\pm$ to $x_\mp$ produces an average flux of $\mathcal{B}_\pm$ molecules, where we expect $\mathcal{B}_\pm \sim O(\Omega)$. The probability to be in state $x_\pm$ is

$$p_\pm \equiv \frac{r_\mp}{r_- + r_+}. \tag{3.38}$$

We will compute D_B using large deviation theory as introduced in Section 3.3.2. In Appendix 3.C, we compute D_B without relying on large deviation theory.

Combining these two contributions, the tilted operator for this two-system system reads [106]

$$\mathcal{L}(\lambda) = \begin{pmatrix} -r_- + J_-\lambda + D_-\lambda^2 & r_+ e^{\lambda \mathcal{B}_+} \\ r_- e^{\lambda \mathcal{B}_-} & -r_+ + J_+\lambda + D_+\lambda^2 \end{pmatrix}. \tag{3.39}$$

The maximal eigenvalue of $\mathcal{L}(\lambda)$ is

$$\alpha(\lambda) = \mathrm{Tr}\mathcal{L}(\lambda)/2 + \sqrt{(\mathrm{Tr}\mathcal{L}(\lambda))^2/4 - \mathrm{Det}\mathcal{L}(\lambda)}, \tag{3.40}$$

where Tr and Det denote the trace and the determinant, respectively. From Eq. (3.22), the average flux of B is given by

$$\begin{aligned} J_B &= \left.\frac{\partial \alpha(\lambda)}{\partial \lambda}\right|_{\lambda=0} \\ &= p_- J_- + p_+ J_+ + r_- p_- (\mathcal{B}_- + \mathcal{B}_+) \\ &= p_- J_- + p_+ J_+ + O\left(e^{-\Omega|\Delta\phi|}\right) \end{aligned} \tag{3.41}$$

where $\Delta\phi_0 = \phi_0(x_0) - \phi_0(x_-)$. The diffusion coefficient reads

$$\begin{aligned} D_B &= \left.\frac{1}{2}\frac{\partial^2 \alpha(\lambda)}{\partial \lambda^2}\right|_{\lambda=0} \\ &= p_- p_+ \frac{(J_- - J_+)^2}{r_- + r_+} + p_- D_- + p_+ D_+ \\ &+ p_- p_+ (\mathcal{B}_- + \mathcal{B}_+)(p_+ - p_-)(J_- - J_+) \\ &+ \frac{1}{2}(\mathcal{B}_- + \mathcal{B}_+)^2 p_- p_+ (r_- p_+ + p_- r_+) \\ &= p_- p_+ \frac{(J_- - J_+)^2}{r_- + r_+} + O(\Omega). \end{aligned} \tag{3.42}$$

Method	Exponential prefactor δ ($D_\sigma \propto \mathrm{e}^{\delta\Omega}$)
Eq. (3.34)	0.0846 ± 0.0005
Eq. (3.43)	0.0846 ± 0.0005
Eq. (3.74)	0.0845 ± 0.0006
Eq. (3.44)	0.0823

Table 3.1: Scaling of the diffusion coefficient D_σ obtained with the CME, Eq. (3.34), the two-state model, Eq. (3.43), the Fokker-Planck equation, Eq. (3.74) and by evaluating the height of the effective potential barrier separating the two fixed-points, Eq. (3.44). The maximum logarithm of the diffusion coefficient is fitted with an linear function of the system size Ω and the errors is given with 95% confidence bounds.

We insert the transition rates Eq. (3.37) into the previous expression and obtain

$$D_B\big|_{\Delta\mu^{\mathrm{bi}}} = p_- p_+ \left(J_- - J_+\right)^2 \frac{\mathrm{e}^{\Omega[\phi_0(x_0)-\phi_0(x_\pm)]}\pi\Omega}{f(x_-)\sqrt{-\phi_0''(x_0)\phi_0''(x_\pm)}}, \tag{3.43}$$

where $f(x_-) = f(x_+)$ and $\phi_0(x_-) = \phi_0(x_+)$ at the bistable point. As a main result, we have thus shown that the diffusion coefficient scales as

$$D_B\big|_{\Delta\mu^{\mathrm{bi}}} \propto \mathrm{e}^{\Omega[\phi_0(x_0)-\phi_0(x_\pm)]}, \tag{3.44}$$

where the exponential prefactor is the height of the effective potential barrier between the two fixed-points. The mean rate of entropy production σ as well as its associated diffusion coefficient D_σ can be evaluated using Eqs. (3.6) and (3.9).

Here, we have assumed that the average flux of B molecules $\mathcal{B}_\pm$ produced during a jump from $x_\pm$ to $x_\mp$ is known and inserted these values in Eq. (3.39). In fact, a trajectory from x_- to x_+ will produce a path-dependent flux of $\mathcal{B}_\pm$. When we perform a coarse-graining of the CME into a two-state model, we lose this information. Nevertheless, as we are only interested in the leading terms of D_B, we have shown that the contributions from jumps between the fixed-points $\mathcal{B}_\pm$ can be neglected close to the bistable point for large system sizes.

3.4.3 Numerical results

We compare the analytical results with numerical evaluations the CME. In Fig. 3.2(b), we show D_σ, Eq. (3.43), and compare it with the numerical results from the CME. We find that D_σ evaluated for the two-state model almost matches the CME close to the bistable point. In Fig. 3.2(c), we show that the diffusion coefficient has an exponential volume-dependence at the bistable point. In Table 3.1, we compare the scaling of the maximum

of the diffusion coefficient obtained numerically and by evaluating Eq. (3.44). The difference between the numerical prefactors and our analytical expression, Eqs. (3.43) and (3.44), is due to finite-size effects.

3.5 Fokker-Planck equation

3.5.1 Stationary solution and diffusion dilemma

The Fokker-Planck equation is usually obtained by truncating the Kramers-Moyal expansion of the chemical master Eq. (3.10). Retaining only the first two terms leads to

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}A_1(x)p(x,t) + \frac{1}{2\Omega}\frac{\partial^2}{\partial x^2}A_2(x)p(x,t) \tag{3.45}$$

where

$$A_1(x) = f(x) - g(x) \quad \text{and} \quad A_2(x) = f(x) + g(x) \tag{3.46}$$

are the first two Kramers-Moyal moments. The stationary distribution of this Fokker-Planck equation is

$$p(x) \propto \exp\Big[-\Omega\Psi_0(x)\Big], \tag{3.47}$$

where

$$\Psi_0(x) = -2\int_0^x \mathrm{d}y\,\frac{A_1(y)}{A_2(y)}. \tag{3.48}$$

However, this potential $\Psi_0(x)$ differs from the CME potential $\phi_0(x)$ in leading Ω orders. As discussed in Section 2.2.3, it is not a suitable choice to describe the transitions between the fixed-points, which are rare events for large system sizes. Hänggi *et al.* [87] proposed a different diffusion process described by a Fokker-Planck equation of the form

$$\frac{\partial}{\partial t}p(x,t) = -\frac{\partial}{\partial x}\Big[L(x)\chi_0(x)p(x,t)\Big] + \frac{1}{\Omega}\frac{\partial^2}{\partial x^2}\Big[L(x)p(x,t)\Big] \equiv \mathcal{L}_0^{\mathrm{FP}}p(x,t), \tag{3.49}$$

where we define, using Eq. (3.15), the generalized thermodynamic force

$$\chi_0(x) \equiv -\frac{\partial\phi_0(x)}{\partial x} \tag{3.50}$$

and an effective diffusion coefficient expressed with a transport coefficient [107, 108]

$$L(x) \equiv \frac{f(x) - g(x)}{\ln f(x) - \ln g(x)}. \tag{3.51}$$

The stationary solution of Eq. (3.49) is

$$p(x) \propto \exp\left[-\Omega\phi_0(x)\right], \tag{3.52}$$

which is equivalent to Eq. (3.14) in leading Ω orders. $\phi_0(x)$ is the large deviation rate function of p(x) [86, 102, 103]. Thus, in the thermodynamic limit ($\Omega \to \infty$), Eq. (3.49) has the same large deviation function as the CME. In other words, Eq. (3.49) will reproduce the correct long-term asymptotic dynamics. However, Eq. (3.49) will not correctly reproduce the stochastic dynamics on the shorter timescale. As shown in Section 2.2.3, the correct diffusion coefficient on the shorter timescale is given by the second Kramers-Moyal moment, which differs from Eq. (3.51) at leading order in Ω. Therefore, there is dilemma when choosing the diffusion term in the Fokker-Planck equation [109, 110]. It is not possible to approximate the CME with a diffusion process that describes correctly the short-time stochastic dynamics and the long-time global dynamics. As shown in Section 3.4, the leading terms of the diffusion coefficient D_B are due to transitions between the fixed-points, thus, Eq. (3.49) is the appropriate Fokker-Planck equation to characterize D_B close to the bistable point.

3.5.2 Behavior of the entropy production at the phase transition

In Section 3.3.2, we derived the tilted operator $\mathcal{L}(\lambda)$ for the CME, Eq. (3.29). Analogously, one can construct a tilted operator $\mathcal{L}(\lambda)$ for the Fokker-Planck equation, see [103, 111] for additional details. For the time-integrated current Z_B, the tiled operator is given by

$$\begin{aligned}\mathcal{L}^{\mathrm{KM}}(\lambda) = &-\left(\frac{\partial}{\partial x} + \lambda\Omega\zeta_B\right) A_1(x) \\ &+ \frac{1}{2\Omega}\left(\frac{\partial}{\partial x} + \lambda\Omega\zeta_B\right)^2 A_2(x)\end{aligned} \tag{3.53}$$

where ζ_B is an operator that only selects transitions in the B channel occurs. For simplicity, we first consider the Kramer-Moyal Fokker-Planck equation. ζ_B acts on the Fokker-Planck coefficients as follows:

$$\begin{aligned}\zeta_B A_1(x) &\equiv \left(\beta^-(x) - \beta^+(x)\right), \\ \zeta_B A_2(x) &\equiv \left(\beta^+(x) + \beta^-(x)\right).\end{aligned} \tag{3.54}$$

The tilted operator coefficient $\mathcal{L}_0^{\mathrm{KM}}$ is given by Eq. (3.45),

$$\mathcal{L}_1^{\mathrm{KM}} \equiv \Omega\left[\beta^-(x) - \beta^+(x)\right] + \frac{\partial}{\partial x}\left(\beta^+(x) + \beta^-(x)\right) \tag{3.55}$$

and

$$\mathcal{L}_2^{\mathrm{KM}} = \frac{\Omega}{2}\left[\beta^+(x) + \beta^-(x)\right]. \tag{3.56}$$

In Appendix 3.D, we derive Eqs. (3.55) and (3.56) by discretizing the tilted operator of the CME.

We now consider the Fokker-Planck Eq. (3.49), where the tilted operator reads

$$\begin{aligned}\mathcal{L}^{\mathrm{FP}}(\lambda) = &-\left(\frac{\partial}{\partial x} + \lambda\Omega\zeta_B\right) L(x)\chi_0(x) \\ &+ \frac{1}{2\Omega}\left(\frac{\partial}{\partial x} + \lambda\Omega\zeta_B\right)^2 2L(x)\end{aligned} \tag{3.57}$$

In order to write the first-order coefficient $\mathcal{L}_1^{\mathrm{FP}}$, we must first compute how the operator ζ_B acts on the diffusion coefficient, $2L(x)$. The drift term reads

$$A_1(x) = L(x)\chi_0(x). \tag{3.58}$$

This expression can also be expressd as [112]

$$A_1(x) = \frac{1}{2P(x)}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n!\,\Omega^n}\frac{\partial^n}{\partial x^n}A_{n+1}(x)P(x) \tag{3.59}$$

where the Kramers-Moyal moments are given by

$$A_n(x) = f(x) + (-1)^n\, g(x) \tag{3.60}$$

and

$$P(x) = \Omega p(\Omega x) \tag{3.61}$$

By expanding the right hand side of Eq. (3.59) in powers of Ω^{-1}, we can express the transport coefficient $L(x)$ as a sum of the Kramers-Moyal moments using Eq. (3.58) leading to

$$L(x) = \frac{1}{2}\sum_{n=0}^{\infty}\frac{1}{(n+1)!}A_{n+2}(x)\left(\chi_0(x)\right)^n. \tag{3.62}$$

Using Eq. (3.54), we can compute how the operator ζ_B acts on $L(x)$. We define the transport coefficient for the B channel as

$$\begin{aligned}L_B(x) &\equiv \zeta_B L(x), \\ &= \frac{1}{2}\sum_{n=0}^{\infty}\frac{1}{(n+1)!}\left(\beta^+(x) + (-1)^n\,\beta^-(x)\right)\left(\chi_0(x)\right)^n, \\ &= \frac{L(x)}{2}\frac{\left(\beta^-(x)f(x) + \beta^+(x)g(x)\right)}{f(x)g(x)}.\end{aligned} \tag{3.63}$$

The tilted operator coefficient $\mathcal{L}_0^{\mathrm{FP}}$ is given by Eq. (3.45),

$$\mathcal{L}_1^{\mathrm{FP}} \equiv \Omega\left[\beta^-(x) - \beta^+(x)\right] + 2\frac{\partial}{\partial x}L_B(x) \tag{3.64}$$

and

$$\mathcal{L}_2^{\mathrm{FP}} \equiv \Omega L_B(x). \tag{3.65}$$

We can write the diffusion coefficient, defined for the CME in Eq. (3.34), as

$$D_B^{\mathrm{FP}} = \int_0^\infty \mathrm{d}x \left\{ \left[\beta^-(x) - \beta^+(x)\right] q_1(x) + \frac{1}{2}\left[\beta^-(x) - \beta^+(x)\right] p(x) \right\} \tag{3.66}$$

where the function $q_1(x)$ is obtained by solving the differential equation

$$\mathcal{L}_0^{\mathrm{FP}} q_1(x) = \left(J_B - \mathcal{L}_1^{\mathrm{FP}}\right) p(x). \tag{3.67}$$

Writing the solution as $q_1(x) = C(x)p(x)$, it can be shown (e.g. using Mathematica) that the solution is

$$\frac{\partial C(x)}{\partial x} = \frac{C_1 + \int_0^x \mathrm{d}y\, p(y)L(y)F(y)}{p(x)L(x)} \tag{3.68}$$

where $C_1 = 0$ due to the boundary condition ($x = 0$) and

$$F(y) = \left\{ \Omega \left[\frac{J_B f(x)g(x) + \beta^+(x)g(x)^2 - \beta^-(x)f(x)^2}{f(y)g(y)L(y)} \right] - \frac{\partial}{\partial y}\left[\frac{\beta^+(y)}{f(y)} + \frac{\beta^-(y)}{g(y)} \right] \right\}. \tag{3.69}$$

Close to the bistable point, the system has three fixed-points, which we order as follows: $0 < x_- < x_0 < x_+$, where $x_\pm$ are stable (i.e. $f'(x_\pm) < g'(x_\pm)$) and x_0 is unstable (i.e. $f'(x_0) > g'(x_0)$). In the limit of large system sizes, we can write $p(x)$ as a sum of two stationary Gaussian processes, Eq. (3.35). Using a saddle-point approximation, Eq. (3.68) becomes

$$\begin{aligned} \frac{\partial C(x)}{\partial x} \approx & \left[\sum_{x_* = (x_-, x_+)} \Theta\left(x - x_*\right) F(x_*) \mathrm{e}^{-\Omega\phi_0(x_*)} \sqrt{\frac{2\pi}{\left|\phi_0''(x_*)\right|\Omega}} \right] \\ & \mathrm{e}^{\Omega\phi_0(x_0)} \sqrt{\frac{2\pi}{\left|\phi_0''(x_0)\right|\Omega}} \exp\left[\frac{-\Omega\phi_0''(x_0)\left(x - x_0\right)^2}{2} \right]. \end{aligned} \tag{3.70}$$

where $\Theta(x)$ is the Heaviside function. We can integrate this expression and obtain

$$C(x) \approx \Theta\left(x - x_-\right) \mathrm{e}^{\Omega[\phi_0(x_0) - \phi_0(x_-)]} \frac{2\pi F(x_-)}{\Omega\sqrt{\left|\phi_0''(x_-)\phi_0''(x_0)\right|}} + C_0, \tag{3.71}$$

where C_0 is a normalization constant. By multiplying Eq. (3.71) with Eq. (3.35), we obtain

$$q_1(x) \approx \frac{1}{\mathcal{Z}_G} \frac{2\pi F(x_-) \mathrm{e}^{\Omega[\phi_0(x_0) - \phi_0(x_-) - \phi_0(x_+)]}}{\Omega f(x_+)\sqrt{\left|\phi_0''(x_-)\phi_0''(x_0)\right|}} \exp\left[\frac{-\Omega\phi_0''(x_+)\left(x - x_+\right)^2}{2} \right] + C_0 p(x). \tag{3.72}$$

where the normalization C_0 is chosen such that

$$\int_0^\infty \mathrm{d}x\, q_1(x) = 0. \tag{3.73}$$

Close to the bistable point, we show that the diffusion coefficient diverges as an exponential function of the system size. We can neglect the term $\int_0^\infty \mathrm{d}x\, \mathcal{L}_2 p(x)$ as it is of order 1 as shown in Section 3.4.2. Eq. (3.66) then becomes

$$\begin{aligned} D_B^{FP} &\approx \int_0^\infty \mathrm{d}x \left[\beta^-(x) - \beta^+(x) - J_B\right] q_1(x) \\ &\approx \frac{1}{\mathcal{Z}_G}\left(\frac{2\pi}{\Omega}\right)^{3/2} \frac{F(x_-)\mathrm{e}^{\Omega[\phi_0(x_0)-\phi_0(x_-)-\phi_0(x_+)]}}{f(x_+)\sqrt{\left|\phi_0''(x_-)\phi_0''(x_0)\phi_0''(x_+)\right|}} \left[\beta^-(x_+) - \beta^+(x_+) - J_B\right]. \end{aligned} \tag{3.74}$$

When the system is bistable, the diffusion coefficient diverges as an exponential law of the system size, namely,

$$D_B^{FP} \propto \mathrm{e}^{\Omega\left[\phi_0(x_0)-\phi_0(x_-)-\phi_0(x_+)-\Omega^{-1}\ln \mathcal{Z}_G\right]}, \tag{3.75}$$

where the normalization constant $\mathcal{Z}_G$ is given by Eq. (3.36). At the bistable point $\Delta\mu^{\mathrm{bi}}$, both stable fixed-points have the same potential height [113], i.e. $\phi_0(x_-) = \phi_0(x_+)$. The fourth term in Eq. (3.75) reduces to

$$\lim_{\Omega\to\infty} \Omega^{-1} \ln \mathcal{Z}_G\Big|_{\Delta\mu^{\mathrm{bi}}} = -\phi_0(x_-). \tag{3.76}$$

By inserting this expression into Eq. (3.75), we show that the diffusion coefficient scales as

$$D_B^{FP}\big|_{\Delta\mu^{\mathrm{bi}}} \propto \mathrm{e}^{\Omega[\phi_0(x_0)-\phi_0(x_-)]}, \tag{3.77}$$

where the exponential prefactor is the same as in Eq. (3.44).

3.5.3 Numerical results

In Fig. 3.2, we plot the diffusion coefficient obtained by evaluating Eq. (3.74) with the parameters given in Section 3.3.3. Close to the bistable point, Eq. (3.74) matches the results obtained for the two-state model, Eq. (3.43) Below the bistable point, Eq. (3.74) deviates from the CME and two-state model expressions. This could be explained by the use of saddle-point approximations, which are only valid when the distribution $p(x)$ has two "distinct" peaks.

3.6 Conclusion

We have investigated the fluctuations of the entropy production at the phase transition occurring in a paradigmatic model of biochemical switches. A control parameter for this phase transition is the thermodynamic force driving the system out of equilibrium. The mean entropy production rate has a discontinuity with respect to the thermodynamic force at the phase transition and fluctuations, which are quantified by the diffusion coefficient that diverges. First, we have computed the diffusion coefficient numerically for the chemical master equation. Second, we have derived an analytical expression of the diffusion coefficient for an effective two-state model. We find that the diffusion coefficient from the two-state model slightly underestimates the diffusion coefficient from the chemical master equation. This difference could be explained by the coarse-graining procedure, which is known to underestimate fluctuations far from equilibrium [114]. Finally, we have shown that the diffusion coefficient has an exponential volume-dependence at the bistable point, where the exponential prefactor is given by the height of the effective potential barrier between the two fixed-points.

In this chapter, we have considered Schlögl's model as a simple model for a nonequilibrium first-order phase transitions. We expect that models with additional chemical reactions or species show qualitatively the same behavior at the phase transition. For bistable systems with multiple species, one can introduce reaction coordinates along which the system becomes effectively one-dimensional. More generally, we expect that diffusion coefficients associated with currents or the entropy production can be computed at first-order phase transitions for a large class of nonequilibrium systems by describing them with discrete jump processes. The exponential volume dependence discussed here should then be generic for these cases.

Appendices to chapter 3

3.A Relation between the cumulants of currents in a system with two reaction channels

Here, we will prove that time-integrated currents $Z_A(t)$ and $Z_B(t)$, which are defined in Eq. (3.4), have the same cumulants. We will rely on the large deviation theory [101–103], which is introduced in Section 3.3.2.

The tilted operators are defined for general observables in Eqs. (3.25) and (3.29). For the A reaction channel, it reads

$$\begin{aligned}(\mathcal{L}_A(\lambda))_{i,j} \equiv{}& \delta_{i,j+1}\left(\alpha_i^+ \mathrm{e}^{\lambda} + \beta_i^+\right) + \delta_{i,j-1}\left(\alpha_i^- \mathrm{e}^{-\lambda} + \beta_i^-\right) \\ &- \delta_{i,j}\left(\alpha_i^+ + \alpha_i^- + \beta_i^+ + \beta_i^+\right)\end{aligned} \tag{3.78}$$

and for the B reaction channel,

$$\begin{aligned}(\mathcal{L}_B(\lambda))_{i,j} \equiv{}& \delta_{i,j+1}\left(\alpha_i^+ + \beta_i^+ \mathrm{e}^{-\lambda}\right) + \delta_{i,j-1}\left(\alpha_i^- + \beta_i^- \mathrm{e}^{\lambda}\right) \\ &- \delta_{i,j}\left(\alpha_i^+ + \alpha_i^- + \beta_i^+ + \beta_i^+\right),\end{aligned} \tag{3.79}$$

where $\alpha_i^{\pm}$ and $\beta_i^{\pm}$ are the transition rates for the A and B channels, respectively. A simple calculation shows that $L_A(\lambda)$ and $L_B(\lambda)$ are related by the following symmetry

$$\mathcal{L}_B(\lambda) = \mathcal{A}^{-1}\mathcal{L}_A(\lambda)\mathcal{A}, \tag{3.80}$$

where

$$\mathcal{A}_{i,j} = \delta_{i,j}\mathrm{e}^{\lambda j}. \tag{3.81}$$

As Eq. (3.80) describes a similarity transformation, $\mathcal{L}_A(\lambda)$ and $\mathcal{L}_B(\lambda)$ have the same eigenvalues. It then follows that $Z_A(t)$ and $Z_B(t)$ have the same scaled cumulant generating function, Eq. (3.21), as it is given by the largest eigenvalue of the tilted operator [115].

3.B Calculation of the coarse-grained rates for the two-state model

Here, we derive the transition rates for the two-state model introduced in Section 3.4. Hänggi *et al.* show that the rates between the fixed-points depend exponentially on

the system size [87]. Moreover, they show that the two-term truncated Kramers-Moyal equation does not lead to the correct expression and consider Fokker-Planck equation with an effective diffusion term that leads to the same stationary distribution as the CME. In this appendix, we will follow a different method proposed by Hinch and Chapmann [88].

In the limit $\Omega \to \infty$, we calculate an approximation solution of the master equation, Eq. (3.10), by introducing the WKB ansatz

$$\hat{p}^{\Omega}(x) \approx R(x,\Omega)\mathrm{e}^{-\Omega\phi_0(x)}, \tag{3.82}$$

with

$$R(x,\Omega) \equiv \sum_{k=0}^{\infty} \Omega^{-k} R_k(x). \tag{3.83}$$

We expand the exponential term

$$\mathrm{e}^{-\Omega\phi_0(x\pm\Omega^{-1})} = \mathrm{e}^{-\Omega\phi_0(x)}\mathrm{e}^{\mp\phi_0'(x)}\left(1+\Omega^{-1}\frac{\phi_0''(x)}{2}+O(\Omega^{-2})\right). \tag{3.84}$$

and the ansatz

$$\hat{p}^{\Omega}(x\pm\Omega^{-1}) = \hat{p}^{\Omega}(x)\mathrm{e}^{\mp\phi_0'(x)}\left(1\mp\Omega^{-1}\frac{\partial}{\partial x}\ln\left(\frac{1}{R(x,\Omega)}\right)+\Omega^{-1}\frac{\phi_0''(x)}{2}\right)+O(\Omega^{-2}) \tag{3.85}$$

By inserting this ansatz into Eq. (3.10), we obtain

$$\begin{aligned} 0 = &\left(f(x)-\Omega^{-1}f'(x)\right)\mathrm{e}^{\phi_0'(x)}\left(1+\Omega^{-1}\frac{\partial}{\partial x}\ln\left(\frac{1}{R(x,\Omega)}\right)+\Omega^{-1}\frac{\phi_0''(x)}{2}\right)\\ &+\left(g(x)-\Omega^{-1}g'(x)\right)\mathrm{e}^{-\phi_0'(x)}\left(1-\Omega^{-1}\frac{\partial}{\partial x}\ln\left(\frac{1}{R(x,\Omega)}\right)+\Omega^{-1}\frac{\phi_0''(x)}{2}\right)\\ &-\left(f(x)+g(x)\right). \end{aligned} \tag{3.86}$$

At leading order in Ω^{-1}, we find

$$0 = f(x)\left(\mathrm{e}^{\phi_0'(x)}-1\right)+g(x)\left(\mathrm{e}^{-\phi_0'(x)}-1\right), \tag{3.87}$$

where the solution is

$$\phi_0(x) = -\int_0^x \mathrm{d}y \ln\left(\frac{f(y)}{g(y)}\right). \tag{3.88}$$

Equating terms at order Ω^{-1} in Eq. (3.86) leads to

$$\frac{R_0'(x)}{R_0(x)} = -\frac{f'(x)}{2f(x)}-\frac{g'(x)}{2g(x)}. \tag{3.89}$$

The solution of this equation is

$$R_0(x) = \frac{A}{\sqrt{f(x)g(x)}}, \tag{3.90}$$

where A is a constant. Using Eqs. (3.88) and (3.90), we obtain the outer (WKB) solution

$$\hat{p}^{\Omega}(x) \approx \frac{A}{\sqrt{f(x)g(x)}} \exp\left[-\Omega \int_0^x \mathrm{d}y \ln\left(\frac{f(y)}{g(y)}\right)\right]. \tag{3.91}$$

At the fixed-points, when $f(x) = g(x)$, the right hand side of Eq. (3.87) vanishes and the outer solution will fail. Close tp the fixed-points, we now compute the inner solution.

In the bistable case there are three fixed-points, $x_* = x_-, x_0, x_+$, which we order as follows: $0 < x_- < x_0 < x_+$, where $x_\pm$ are stable (i.e. $f'(x_\pm) < g'(x_\pm)$) and x_0 is unstable (i.e. $f'(x_0) > g'(x_0)$). Close to the fixed-points, we introduce a local variable

$$x = x_* + \Omega^{-1/2} z \tag{3.92}$$

and denote $P(z)$ the inner solution. We write Eq. (3.10) with this local variable and obtain at leading order in Ω^{-1}

$$0 = \frac{\mathrm{d}^2 P(z)}{\mathrm{d}z^2} + \gamma(x_*)\left(z\frac{\mathrm{d}P(z)}{\mathrm{d}z} + P(z)\right) + O\left(\Omega^{-3/2}\right) \tag{3.93}$$

where we define

$$\gamma(x_*) \equiv \frac{g'(x_*) - f'(x_*)}{f(x_*)}. \tag{3.94}$$

The inner solution of Eq. (3.93) reads

$$P(z) = \mathrm{e}^{\gamma(x_*)z^2/2}\left(B + C\int_0^z \mathrm{d}y\, \mathrm{e}^{\gamma(y)y^2/2}\right), \tag{3.95}$$

where B and C are constant determined by asymptotically matching the outer and inner solution, respectively, Eqs. (3.91) and (3.95), using Van Dyke's rule [116].

Solution at the stable fixed-point

We denote $\hat{p}^{\Omega}_-$ the solution left of the fixed-point ($x < x_*$) with amplitude $A = A_-$ in Eq. (3.91) and $\hat{p}^{\Omega}_+$ the solution right of the fixed-point ($x > x_*$) with amplitude $A = A_+$. The constraint $P(z) > 0$ requires $C = 0$. Matching Eqs. (3.91) and (3.95) asymptotically ($z \to \pm\infty$) leads to

$$B = \frac{A_- \mathrm{e}^{-\Omega\phi_0(x_*)}}{\sqrt{f(x_*)g(x_*)}} \quad \text{and} \quad A_+ = A_-. \tag{3.96}$$

This solution is the stationary Gaussian considered in Section 3.4.

Solution at the unstable fixed-point

We denote $\hat{p}_-^\Omega$ the solution left of the fixed-point ($x < x_*$) with amplitude $A = A_L$ in Eq. (3.91) and $\hat{p}_+^\Omega$ the solution right of the fixed-point ($x > x_*$) with amplitude $A = A_R$. The outer solutions, Eq. (3.91), written with the local variable defined in Eq. (3.92) read in the limit $\Omega \to \infty$

$$\hat{p}_\pm^\Omega = \frac{A_\pm \mathrm{e}^{-\Omega\phi_0(x_0)}}{\sqrt{f(x_0)g(x_0)}} \mathrm{e}^{-\gamma(x_0)z^2/2}. \tag{3.97}$$

Matching these solutions to the inner solutions, Eq. (3.95), asymptotically leads to

$$\begin{aligned} B &= \frac{A_+ \mathrm{e}^{-\Omega\phi_0(x_0)}}{\sqrt{f(x_0)g(x_0)}}, \\ C &= \frac{(A_- - A_+)\mathrm{e}^{-\Omega\phi_0(x_0)}}{\sqrt{f(x_0)g(x_0)}} \sqrt{\frac{-\gamma(x_0)}{2\pi}} \end{aligned} \tag{3.98}$$

We define $P_\pm$ the probability to be, respectively, on left or right side of the unstable point x_0 as

$$P_\pm \equiv A_\pm T_\pm \tag{3.99}$$

where

$$T_- \equiv \Omega \int_0^{x_0} \mathrm{d}x \frac{\mathrm{e}^{-\Omega\phi_0(x)}}{\sqrt{f(x)g(x)}} \quad \text{and} \quad T_+ \equiv \Omega \int_{x_0}^{\infty} \mathrm{d}x \frac{\mathrm{e}^{-\Omega\phi_0(x)}}{\sqrt{f(x)g(x)}}. \tag{3.100}$$

We can approximate these integrals with a saddle-point approximation leading to

$$T_\pm \approx \frac{\mathrm{e}^{-\Omega\phi_0(x_\pm)}}{f(x_\pm)} \sqrt{\frac{2\pi\Omega}{\gamma(x_\pm)}}. \tag{3.101}$$

Using the conservation of probability

$$P_- + P_+ = A_- T_- + A_+ T_+ = 1, \tag{3.102}$$

we can write

$$A_+ = \frac{1 - A_- T_-}{T_+}, \tag{3.103}$$

which leaves one free parameter A_- to be determined.

Second largest eigenvalue λ_1

We want to compute the eigenvalues of the master equation operator $\mathcal{L}_0$, defined in Eq. (3.10). The eigenvalue problem

$$\mathcal{L}_0\psi(x) = \lambda\psi(x) \tag{3.104}$$

has a zero eigenvalue λ_0, where the corresponding eigenvector $\psi^{(0)}$ is given by the stationary distribution Eq. (3.14). Differentiating Eq. (3.91) with respect to A_- leads to an eigenfunction $\psi^{(1)}(x)$ with an exponentially small error [88],

$$\psi^{(1)}(x) \approx \begin{cases} \frac{\mathrm{e}^{-\Omega\phi_0(x)}}{\sqrt{f(x)g(x)}} & \text{for } x < x_0, \\ \frac{\mathrm{e}^{-\Omega\phi_0(x_0)}}{\sqrt{f(x_0)g(x_0)}}\mathrm{e}^{\gamma(x_0)z^2/2}\left[-\frac{T_-}{T_+} + \sqrt{\frac{-\gamma(x_0)}{2\pi}}\left(1+\frac{T_-}{T_+}\right)\int_z^\infty \mathrm{d}y\, \mathrm{e}^{\gamma(x_0)y^2/2}\right] & \text{for } x \approx x_0, \\ -\frac{T_-}{T_+}\frac{\mathrm{e}^{-\Omega\phi_0(x)}}{\sqrt{f(x)g(x)}} & \text{for } x > x_0. \end{cases} \tag{3.105}$$

For simplicity, we consider the transposed eigenvalue problem

$$\mathcal{L}_0^{\mathrm{T}}\widetilde{\psi}(x) = \lambda\widetilde{\psi}(x), \tag{3.106}$$

where the left eigenvector for $\lambda_0 = 0$ is $\widetilde{\psi}^{(0)} = 1$. We write the inner equation using the local variable defined in Eq. (3.92) and obtain at leading order in Ω^{-1}

$$0 = \frac{\mathrm{d}^2\widetilde{\psi}^{(1)}}{\mathrm{d}z^2} - \gamma(x_*)\frac{\mathrm{d}\widetilde{\psi}^{(1)}}{\mathrm{d}z} + O\left(\Omega^{-3/2}\right), \tag{3.107}$$

where the solution is

$$\widetilde{\psi}^{(1)}(z) \approx B + C\int_{-\infty}^{z} \mathrm{d}y\, \mathrm{e}^{\gamma(y)y^2/2}. \tag{3.108}$$

At the stable fixed-points, the outer solution sets $C = 0$ so $\widetilde{\psi}^{(1)}(z)$ remains constant. Thus, matching the outer and inner solution leads to

$$\widetilde{\psi}^{(1)}(x) \approx \begin{cases} B_- & \text{for } x < x_0, \\ B_- + (B_+ - B_-)\sqrt{\frac{\gamma(x_0)}{2\pi}}\int_{-\infty}^{z} \mathrm{d}y\, \mathrm{e}^{\gamma(x_0)y^2/2} & \text{for } x \approx x_0, \\ B_+ & \text{for } x > x_0. \end{cases} \tag{3.109}$$

The constant $B\pm$ can be determined using the orthonormality condition

$$\int_0^\infty \mathrm{d}x\, \widetilde{\psi}^{(i)}(x)\psi^{(j)}(x) = \delta_{i,j}\,, \tag{3.110}$$

leading to

$$\widetilde{\psi}^{(1)}(z) \approx \begin{cases} \frac{T_+}{T_-(T_-+T_+)} & \text{for } x < x_0, \\ \frac{T_+}{T_-(T_-+T_+)} + \frac{1}{T_-}\sqrt{\frac{\gamma(x_0)}{2\pi}} \int_{-\infty}^{z} \mathrm{d}y\, \mathrm{e}^{\gamma(x_0)y^2/2} & \text{for } x \approx x_0, \\ -\frac{1}{(T_-+T_+)} & \text{for } x > x_0. \end{cases} \tag{3.111}$$

We can compute the eigenvalue λ_1. We write the inner-product, Eq. (3.104), with a suitable test function H as

$$\int_0^\infty \mathrm{d}y\, H(y)\mathcal{L}_0\psi^{(1)}(y) = \lambda_1 \int_0^\infty \mathrm{d}y\, H(y)\psi^{(1)}(y). \tag{3.112}$$

Here, we choose H such that

$$H \equiv \begin{cases} 1 & \text{for } x \le x_0, \\ 0 & \text{for } x > x_0. \end{cases} \tag{3.113}$$

Writing Eq. (3.112) as a discrete sum, we obtain

$$\sum_{n=0}^{x_0\Omega} \mathcal{L}\psi^{(1)}(n/\Omega) = -f(x_0)\psi^{(1)}(x_0) + g(x_0+\Omega^{-1})\psi^{(1)}(x_0+\Omega^{-1}). \tag{3.114}$$

By evaluating $\psi^{(1)}(x)$ using Eq. (3.105), this expression becomes

$$\begin{aligned} \sum_{n=0}^{x_0\Omega} \mathcal{L}\psi^{(1)}(n/\Omega) &= f(x_0)\left[\psi^{(1)}(z=\Omega^{-1}) - \psi^{(1)}(z=0)\right] \\ &\approx -\mathrm{e}^{\Omega\phi_0(x_0)}\sqrt{-\gamma(x_0)}2\pi\Omega\left(1+\frac{T_-}{T_+}\right). \end{aligned} \tag{3.115}$$

The right hand side of Eq. (3.112) can be evaluated using Eqs. (3.100) and (3.101)

$$\lambda_1 \int_0^{x_0} \mathrm{d}y\psi^{(1)}(y) = \lambda_1 T_- \approx \lambda_1 \frac{\mathrm{e}^{-\Omega\psi_0(x_-)}}{f(x_-)} \frac{2\pi\Omega}{\gamma(x_-)}. \tag{3.116}$$

Combining Eqs. (3.115) and (3.116), we obtain the eigenvalue

$$\lambda_1 \approx -\frac{\mathrm{e}^{\Omega[\phi_0(x_0)-\phi_0(x_-)]} f(x_-)\sqrt{-\gamma(x_0)\gamma(x_-)}}{2\pi\Omega}\left(1+\frac{T_-}{T_+}\right). \tag{3.117}$$

Transition rates between the stable fixed-points

Following [117], we look for a quasi-steady state linearized solution of the form

$$p(x,t) \approx \hat{p}^{\Omega}\left(x, A_{-}(t)\right) + r(x,t), \tag{3.118}$$

where we expect $r(x,t) \ll \hat{p}^{\Omega}(x,t)$ in the long time limit. By inserting this solution into Eq. (3.10), we obtain

$$\frac{\partial r(x,t)}{\partial t} = \mathcal{L}_0 r(x,t) - \frac{\partial}{\partial t}\hat{p}^{\Omega}\left(x, A_{-}(t)\right) + \mathcal{L}_0 \hat{p}^{\Omega}\left(x, A_{-}(t)\right). \tag{3.119}$$

We write $r(x,t)$ with the eigenfunctions $\psi^{(i)}(x)$ defined Eq. (3.104) as

$$r(x,t) = \sum_{i=0}^{\infty} c_i(t)\psi^{(i)}(x) \tag{3.120}$$

Using the orthonormality condition, Eq. (3.110), we project Eq. (3.119) onto $\widetilde{\psi}^{(1)}(x)$ leading to the differential equation

$$c_1'(t) - \lambda_1 c_1(t) = \int_0^{\infty} \mathrm{d}x\, \widetilde{\psi}^{(1)}(x)\mathcal{L}_0\hat{p}^{\Omega}\left(x, A_{-}(t)\right) - \int_0^{\infty} \mathrm{d}x\, \widetilde{\psi}^{(1)}(x)\frac{\partial}{\partial t}\hat{p}^{\Omega}\left(x, A_{-}(t)\right) \tag{3.121}$$

As $\lambda_1 < 0$ and is exponentially small, $c_1(t)$ must vanish to ensure that $r(x,t) \ll \hat{p}^{\Omega}(x,t)$ in the long time limit. Thus, the right hand side of Eq. (3.121) must vanish, leading to a differential equation for $A_{-}(t)$

$$\int_0^{\infty} \mathrm{d}x\, \widetilde{\psi}^{(1)}(x)\mathcal{L}_0\hat{p}^{\Omega}\left(x, A_{-}(t)\right) = \int_0^{\infty} \mathrm{d}x\, \widetilde{\psi}^{(1)}(x)\frac{\partial}{\partial t}\hat{p}^{\Omega}\left(x, A_{-}(t)\right). \tag{3.122}$$

The right hand side becomes

$$\int_0^{\infty} \mathrm{d}x\, \widetilde{\psi}^{(1)}(x)\frac{\partial}{\partial t}\hat{p}^{\Omega}\left(x, A_{-}(t)\right) = A_{-}'(t) \tag{3.123}$$

Using Eqs. (3.91), (3.100) and (3.111), the left hand side can be written as

$$\begin{aligned}\int_0^{\infty} \mathrm{d}x\, \widetilde{\psi}^{(1)}(x)\mathcal{L}_0\hat{p}^{\Omega}\left(x, A_{-}(t)\right) &= \int_0^{\infty} \mathrm{d}x\, \mathcal{L}_0^{\mathrm{T}}\widetilde{\psi}^{(1)}(x)\hat{p}^{\Omega}\left(x, A_{-}(t)\right) \\ &= \lambda_1 \int_0^{\infty} \mathrm{d}x\, \widetilde{\psi}^{(1)}(x)\hat{p}^{\Omega}\left(x, A_{-}(t)\right) \\ &\approx \lambda_1 \left(\widetilde{\psi}^{(1)}(x)A_{(}t)\right),\end{aligned} \tag{3.124}$$

where we use a saddle-point approximation in the last line. By replacing A_+ with Eq. (3.103), Eq. (3.122) becomes a differential equation for A_-,

$$A'_-(t) = |\lambda_1| \left(\frac{1}{T_- + T_+} - A_- \right). \tag{3.125}$$

Finally, we can compute the coarse-grained transition rates of the two-state model considered in Section 3.4. The rate of transfer of probability of a distribution initially in the left or right side of the unstable fixed-point, respectively, with $A_+ = 0$ and $A_+ = 0$. Using Eqs. (3.100), (3.117) and (3.123), the time evolution of the P_- with the initial condition $A_+ = 0$ is

$$P'_-(t) = T_- A'_- = -\frac{e^{\Omega[\phi_0(x_0) - \phi_0(x_-)]} f(x_-) \sqrt{-\gamma(x_0)\gamma(x_-)}}{2\pi\Omega} P_-(t). \tag{3.126}$$

Likewise, one can show that the time evolution of the P_+ with the initial condition $A_- = 0$ is

$$P'_+(t) = T_+ A'_+ = -\frac{e^{\Omega[\phi_0(x_0) - \phi_0(x_+)]} f(x_+) \sqrt{-\gamma(x_0)\gamma(x_+)}}{2\pi\Omega} P_+(t). \tag{3.127}$$

These two expressions corresponds to Eq. (3.37) in Chapter 3.

3.C Calculation of the diffusion coefficient without relying on large deviation theory

Here, we present a derivation of the diffusion coefficient without relying on large deviation theory. We consider the two-state model introduced in Section 3.4. For simplicity, we neglect contributions from jumps between fixed-points $\mathcal{B}_\pm$ and the diffusion around the fixed-points, see 3.4.2 for further explanations.

Along a stochastic trajectory $n(t)$, the time-integrated current of B molecules is given by

$$Z_B = \int_0^T \mathrm{d}t \left(J_- \delta_{n(t),-} + J_+ \delta_{n(t),+} \right). \tag{3.128}$$

where $J_\pm$ is the flux of B molecules in state $x_\pm$. The average flux is

$$J_B = \lim_{T\to\infty} \frac{\langle Z_B \rangle}{T} = \sum_{n=-,+} p_n J_n. \tag{3.129}$$

To compute the diffusion coefficient, we will consider a shifted system where the flux is 0 in state x_- and $(J_+ - J_-)$ in state x_+. The shifted time-integrated current is

$$\widetilde{Z}_B \equiv \int_0^T \mathrm{d}t \, (J_+ - J_-) \, \delta_{n(t),+}, \tag{3.130}$$

and its associated flux

$$\langle \widetilde{Z}_B \rangle = \langle Z_B \rangle - J_- = T\,(J_+ - J_-)\,p_+. \tag{3.131}$$

The second moment of $\widetilde{Z}_B$ is given by

$$\langle \widetilde{Z}_B^2 \rangle = (J_+ - J_-)^2 \int_0^T \mathrm{d}t \int_0^T \mathrm{d}t' \underbrace{\langle \delta_{n(t),+} \delta_{n(t'),+} \rangle}_{p(+,t;+,t')} \tag{3.132}$$

where $p(+,t;+,t')$ is the joint probability to be in state x_+ at times t and t'. In the steady-state it is equal to $p(+,t-t';+,0)$. We solve the two-state master equation and obtain

$$\begin{aligned} p(+,\tau;+,0) &= p_+\, p(+,\tau|+,0) \\ &= p_+ \left(p_+ - (p_+ - 1)\mathrm{e}^{-(r_- + r_+)\tau} \right). \end{aligned} \tag{3.133}$$

By inserting this expression into Eq. (3.132), we can compute the variance which does not depend on the shift. We obtain

$$\begin{aligned} \langle Z_B^2 \rangle - \langle Z_B \rangle^2 &= \langle \widetilde{Z}_B^2 \rangle - \langle \widetilde{Z}_B \rangle^2 \\ &= 2p_+\,(J_+ - J_-)^2 \int_0^T \mathrm{d}t \int_0^t \mathrm{d}\tau \left(p_+ - (p_+ - 1)\mathrm{e}^{-(r_- + r_+)\tau} \right) - \langle \widetilde{Z}_B \rangle^2 \\ &= 2p_+\,(J_+ - J_-)^2 \left(\frac{(1-p_+)}{(r_- + r_+)^2} \left(\mathrm{e}^{-(r_- + r_+)T} - 1 \right) + \frac{(1-p_+)\,T}{(r_- + r_+)} \right). \end{aligned} \tag{3.134}$$

Finally, we get the diffusion coefficient

$$D_B = \lim_{T\to\infty} \frac{\langle Z_B^2 \rangle - \langle Z_B \rangle^2}{2T} = p_- p_+ \frac{(J_+ - J_-)^2}{r_- + r_+}. \tag{3.135}$$

3.D Tilted operator for the Fokker-Planck equation

We start by constructing the tilted operator $\mathcal{L}^{\mathrm{KM}}(\lambda)$ for the Fokker-Planck Eq. (3.45) by discretizing the master operator Eq. (3.10). We obtain

$$\begin{aligned} \mathcal{L}_0 = &\left[f_n - \Omega^{-1} f_n' + \frac{\Omega^{-2}}{2} f_n'' \right] \left[p_n - \Omega^{-1} p_n' + \frac{\Omega^{-2}}{2} p_n'' \right] \\ &+ \left[g_n + \Omega^{-1} g_n' + \frac{\Omega^{-2}}{2} g_n'' \right] \left[p_n + \Omega^{-1} p_n' + \frac{\Omega^{-2}}{2} p_n'' \right] \\ &- (f_n + g_n)\, p_n + O(\Omega^{-3}) \end{aligned} \tag{3.136}$$

and recover with the change of variable $x = n\Omega^{-1}$ the Fokker-Planck operator Eq. (3.45)

$$\mathcal{L}_0^{\mathrm{KM}} = -\frac{\partial}{\partial x}\Big[f(x) - g(x)\Big] + \frac{1}{2\Omega}\frac{\partial^2}{\partial x^2}\Big[f(x) + g(x)\Big]. \tag{3.137}$$

The first coefficient of the titled operator, defined in Eq. (3.29), is

$$\begin{aligned}\mathcal{L}_1 =& -\left[\beta_n^+ - \Omega^{-1}\beta_n^{+\prime} + \frac{\Omega^{-2}}{2}\beta_n^{+\prime\prime}\right]\left[p_n - \Omega^{-1}p_n' + \frac{\Omega^{-2}}{2}p_n''\right] \\ &+ \left[\beta_n^- + \Omega^{-1}\beta_n^{-\prime} + \frac{\Omega^{-2}}{2}\beta_n^{-\prime}\right]\left[p_n + \Omega^{-1}p_n' + \frac{\Omega^{-2}}{2}p_n''\right] + O(\Omega^{-3}).\end{aligned} \tag{3.138}$$

We then recover Eq. (3.55), which reads

$$\mathcal{L}_1^{\mathrm{KM}} = \Omega\left[\beta^-(x) - \beta^+(x)\right] + \frac{\partial}{\partial x}\left[\beta^-(x) + \beta^+(x)\right]. \tag{3.139}$$

The second coefficient of the titled operator is

$$\begin{aligned}\mathcal{L}_2 =& \frac{1}{2}\left[\beta_n^+ - \Omega^{-1}\beta_n^{+\prime} + \frac{\Omega^{-2}}{2}\beta_n^{+\prime\prime}\right]\left[p_n - \Omega^{-1}p_n' + \frac{\Omega^{-2}}{2}p_n''\right] \\ &+ \frac{1}{2}\left[\beta_n^- + \Omega^{-1}\beta_n^{-\prime} + \frac{\Omega^{-2}}{2}\beta_n^{-\prime}\right]\left[p_n + \Omega^{-1}p_n' + \frac{\Omega^{-2}}{2}p_n''\right] + O(\Omega^{-3}).\end{aligned} \tag{3.140}$$

We then recover Eq. (3.55), which reads

$$\mathcal{L}_2^{\mathrm{KM}} = \frac{\Omega}{2}\left[\beta^-(x) + \beta^+(x)\right]. \tag{3.141}$$

Thus, Eq. (3.53) matches the expressions derived in this appendix.

4 Second-order phase transitions in biochemical oscillators

4.1 Introduction

Living systems need biochemical oscillators [118] for the timing and control of several key processes such as circadian rhythms [119, 120] and the cell cycle [121]. Oscillations can only set in if the system is out of equilibrium. As a control parameter for this phase transition, we choose the thermodynamic force driving the system far from equilibrium. For biological systems, the thermodynamic force can be the free energy liberated with the hydrolysis of one ATP molecule. Cao *et al.* [122] have shown that this thermodynamic force must be above a certain threshold for the onset of biochemical oscillations. Therefore, a natural question that arises is whether biochemical oscillators display a phase transition. In other words, what kind of non-analytical behavior do physical observables display at this critical thermodynamic force?

In this chapter, we show that a generic phase transition takes place in biochemical oscillators. As in the well-known theory of nonequilibrium phase transitions [35–39], this phase transition is associated with a Hopf bifurcation, i.e., the onset of limit cycle, in the deterministic rate equations. An observable that characterizes the transition is the steady state distribution of the chemical species that oscillates. This distribution becomes bimodal above the critical force. We analyze the critical behavior of the fluctuating entropy production current. The first derivative of its mean with respect to the thermodynamic force is found to be discontinuous at the critical point and its diffusion coefficient is found to diverge there.

Since these biochemical oscillations can occur in systems with a finite number of molecules that lead to relatively large fluctuations, it is natural to study the precision of biochemical oscillations [122–124]. More broadly, the relation between precision and dissipation in biophysics has been intensively investigated [100, 125–136]. We analyze the relation between the Fano factor associated with the thermodynamic current and the number of coherent oscillations. The Fano factor has been analyzed for theoretical studies related to single molecule experiments [137–141] and has been proposed as an observable that can quantify the precision of biochemical oscillators [124, 142]. Interestingly, this Fano factor has a universal lower bound that depends solely on the thermodynamic force, which follows from the thermodynamic uncertainty relation [20, 21, 106, 143].

The number of coherent oscillations, which is the number of periods for which different stochastic realizations remain coherent with each other, is a standard measure of the precision of biochemical oscillators [122, 123, 144–149]. It can be used to identify the onset of biochemical oscillations, i.e., it is zero below the critical force and becomes non-zero above it [122]. We show that the Fano factor does not properly quantify the precision of biochemical oscillations. This observable that diverges at the critical point, due to the divergence of the diffusion coefficient, is shown to be small below the critical point, indicating high precision even if there are no biochemical oscillations.

We consider three different models for biochemical oscillators: the Brusselator [55], the activator-inhibitor model [122], and a model for the oscillations in the phosphorylation level of KaiC [150], which is a protein related to the regulation of the circadian rhythm of cyanobacteria [151].

The chapter is organized as follows. In Sections 4.2 to 4.4 we introduce the three models and analyze their critical behavior. In Section 4.5 we discuss metrics for the precision of biochemical oscillations, with the comparison between the number of coherent oscillations and the Fano factor. We conclude in Section 4.6. Details of the activator-inhibitor model and the KaiC model are provided in Appendix 4.A and Appendix 4.B, respectively.

4.2 Brusselator

4.2.1 Model definition

The Brusselator is a paradigmatic model for biochemical oscillations [55, 152–154]. It consists of two intermediate species X and Y in a volume Ω. The external bath contains two chemical species A and B at fixed concentrations $[A]$ and $[B]$, respectively. The set of chemical reactions is

$$\begin{aligned} A &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X, \\ B &\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} Y, \\ 2X + Y &\underset{k_3}{\overset{k_3}{\rightleftharpoons}} 3X, \end{aligned} \tag{4.1}$$

where k_1, k_{-1}, k_2, k_{-2} and k_3 are transition rates. For convenience we assume that the transition rates for the forward and backward direction in the third reaction are the same. Note that this Brusselator, first considered in [153] is a modified version of the original Brusselator [55, 152].

The system is driven out of equilibrium due to a difference of chemical potential between A and B, which is written as $\Delta\mu \equiv \mu_B - \mu_A$. For example, consider the following

cycle: a Y molecule is created with rate k_2, then a Y molecule is transformed into an X molecule with rate k_3 and, finally, an X molecule is degraded with rate k_{-1}. This cycle leads to the consumption of substrate B and generation of product A. The thermodynamic force associated with this cycle is

$$\Delta\mu \equiv \ln \frac{k_{-1}k_2[B]}{k_{-2}k_1[A]}, \tag{4.2}$$

where the temperature T and Boltzmann's constant k_B are set to $k_BT = 1$. The above relation between the thermodynamic parameter $\Delta\mu$ and the transition rates is known as generalized detailed balance [6].

The state of the system is determined by two variables, the total number of X molecules n_X and the total number of Y molecules n_Y. The time evolution of $P(N_X, N_Y, t)$, the probability to find the system in state (N_X, N_Y) at time t, is governed by the chemical master equation, which reads

$$\begin{aligned}\frac{\partial}{\partial t}P(N_X, N_Y, t) = &\Big\{\Omega k_1[A]\left[\mathcal{E}_X^- - 1\right] + \Omega k_2[B]\left[\mathcal{E}_Y^- - 1\right] \\ &+ k_{-1}\left(\mathcal{E}_X^+ - 1\right)N_X + k_{-2}\left(\mathcal{E}_Y^+ - 1\right)(N_Y + 1) \\ &+ \frac{k_3}{\Omega^2}\left(\mathcal{E}_X^-\mathcal{E}_Y^+ - 1\right)N_X(N_X - 1)N_Y \\ &+ \frac{k_3}{\Omega^2}\left(\mathcal{E}_X^+\mathcal{E}_Y^- - 1\right)N_X(N_X - 1)(N_X - 2)\Big\}P(N_X, N_Y, t),\end{aligned} \tag{4.3}$$

where we define step operators as

$$\begin{aligned}\mathcal{E}_X^\pm P(N_X, N_Y, t) &\equiv P(N_X \pm 1, N_Y, t), \quad \mathcal{E}_X^\pm N_X \equiv (N_X \pm 1), \\ \mathcal{E}_Y^\pm P(N_X, N_Y, t) &\equiv P(N_X, N_Y \pm 1, t), \quad \mathcal{E}_Y^\pm N_Y \equiv (N_Y \pm 1).\end{aligned} \tag{4.4}$$

The system reaches a nonequilibrium steady state with a steady-state distribution written as $P(N_X, N_Y)$. The marginal distribution of N_X that we evaluate in numerical simulations is defined as $P(N_X) \equiv \sum_{N_Y} P(N_X, N_Y)$

From the master Eq. (4.3), we obtain the equations for the time evolution of the densities

$$\begin{aligned}x &\equiv \sum_{N_X,N_Y} N_X P(N_X, N_Y, t)/\Omega, \\ y &\equiv \sum_{N_X,N_Y} N_Y P(N_X, N_Y, t)/\Omega,\end{aligned} \tag{4.5}$$

in the deterministic limit ($\Omega \to \infty$), which read

$$\begin{aligned}\frac{dx}{dt} &= k_1[A] - k_{-1}x + k_3\left(x^2y - x^3\right), \\ \frac{dy}{dt} &= k_2[B] - k_{-2}y - k_3\left(x^2y - x^3\right).\end{aligned} \tag{4.6}$$

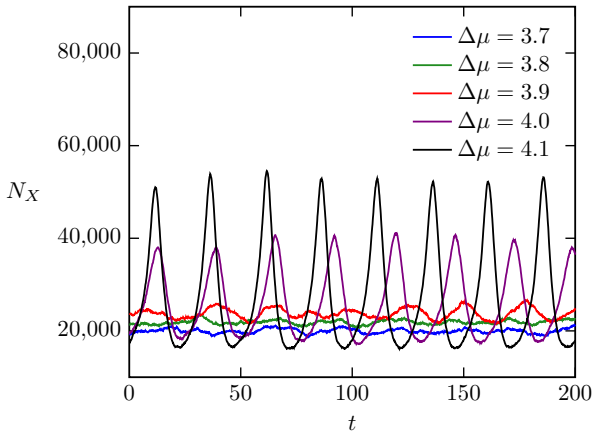

Figure 4.1: Trajectories of N_X at different $\Delta\mu$ in the Brusselator with volume $\Omega = 10^5$, the remaining parameters are given in the main text.

We have performed continuous time Monte Carlo simulations using the Gillespie algorithm [70]. We set the parameters to $k_1 = k_2 = 0.1$, $k_{-1} = k_3 = 1$, $[A] = 1$, $[B] = 3$. The rate k_{-2} is computed with $\Delta\mu$ and the generalized detailed balance relation in Eq. (4.2), where $\Delta\mu$ is the control parameter. In Fig. 4.1, we show stochastic trajectories of this model. For large enough $\Delta\mu$, biochemical oscillations set in.

4.2.2 Results

In the deterministic limit described by Eq. (4.6), for $\Delta\mu \geq \Delta\mu^c \simeq 3.95$ the stationary solution becomes numerically unstable and a limit cycle sets in. The thermodynamic flux associated with this force is the rate of generation of the product A per volume Ω, which must be equal the rate of consumption of B due to the conservation of the number of particles in the reservoir. In the deterministic limit this flux takes the form

$$J_A = k_{-1}x - k_1[A], \tag{4.7}$$

where x is the stationary solution of Eq. (4.6). The rate of entropy production is given by

$$\sigma \equiv J_A \Delta\mu \tag{4.8}$$

in the steady state. As shown in Fig. 4.2(a), there is a discontinuity in the first derivative $\partial\sigma/\partial\Delta\mu$ at the critical point $\Delta\mu^c$ in this deterministic limit.

For a stochastic system, the average thermodynamic flux reads

$$J_A = \Omega^{-1} k_{-1} \sum_{N_X} N_X P(N_X) - k_1[A]. \tag{4.9}$$

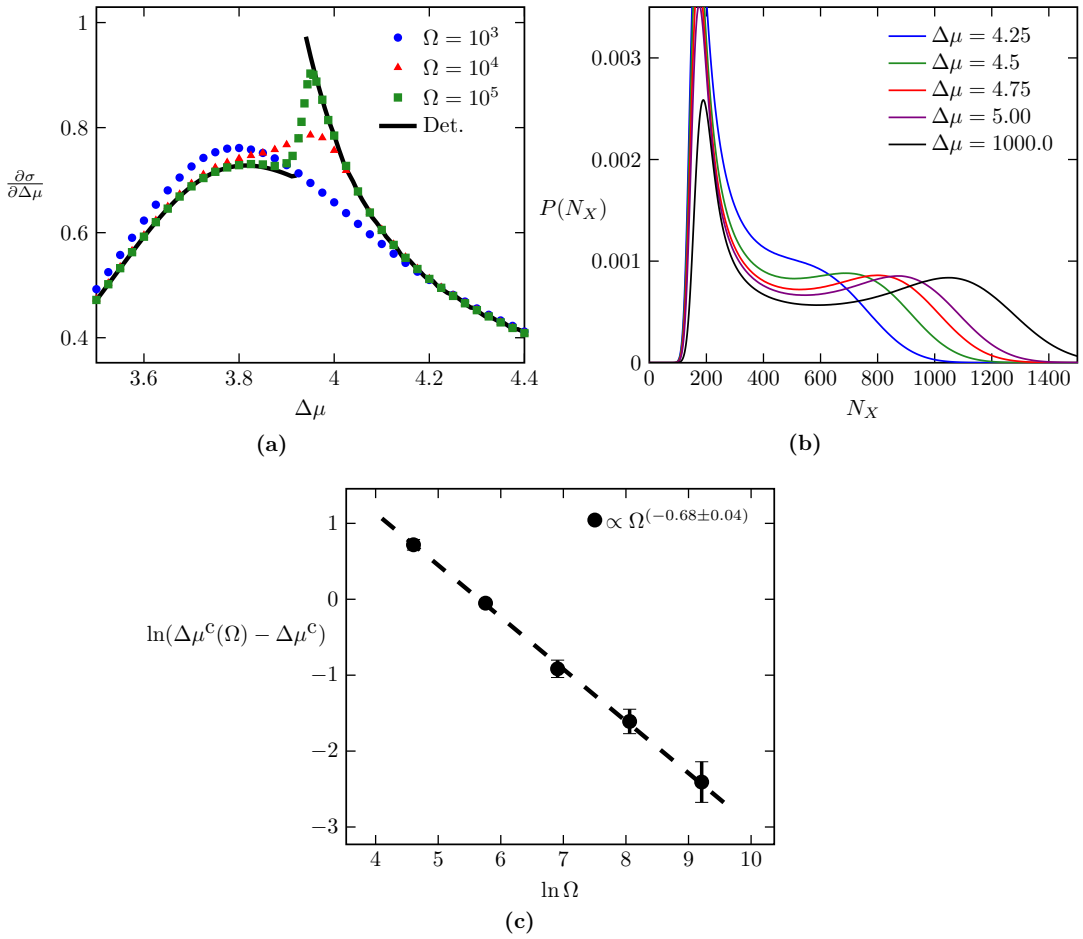

Figure 4.2: Phase transition in the Brusselator. (a) First derivative of the mean entropy production rate σ as a function of $\Delta\mu$. The critical point is $\Delta\mu^c \simeq 3.95$. (b) Stationary distribution of chemical species X for $\Omega = 10^3$ and different values of $\Delta\mu$. (c) Difference between the point at which the distribution in (b) becomes bimodal in a finite system $\Delta\mu^c(\Omega)$ and the critical point $\Delta\mu^c \simeq 3.95$ obtained with the deterministic rate equations.

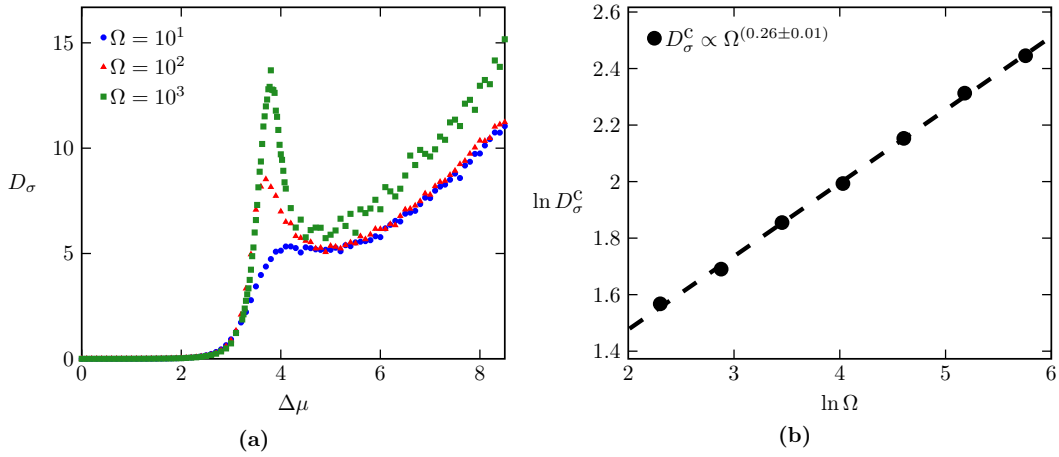

Figure 4.3: Diffusion coefficient for the Brusselator. (a) Diffusion coefficient D_σ as a function of $\Delta\mu$. (b) Maximum diffusion coefficient D_σ^c as a function of Ω.

In Fig. 4.2(a), we plot $\partial\sigma/\partial\Delta\mu$ as a function of $\Delta\mu$. With increasing system size, the curve tends to the deterministic result, with an increase of $\partial\sigma/\partial\Delta\mu$ close to the critical point that gets steeper with increasing volume Ω.

For a finite system, there is a crossover to oscillatory behaviors, as illustrated in Fig. 4.1. The stationary probability distribution $P(N_X)$ also changes at this crossover. Below some finite-size critical point, where biochemical oscillations do not take place, this distribution is unimodal, whereas above this critical point this distribution becomes bimodal. This result is shown in Fig. 4.2(b). We define the finite-size critical point $\Delta\mu^c(\Omega)$ as the minimal $\Delta\mu$ for which the distribution displays a local minimum. In Fig. 4.2(c), we show that $\Delta\mu^c(\Omega)$ converges to $\Delta\mu^c \simeq 3.95$, where the difference $\Delta\mu^c(\Omega) - \Delta\mu^c$ decreases as a power-law with the system size Ω.

Fluctuations related to the thermodynamic flux can be analyzed by considering a stochastic time-integrated current $Z_A(t)$, which is extensive in time. In a stochastic trajectory, this random variable increases by one if an A is produced, which happens if the transition with rate k_{-1} takes place, and it decreases by one if an A is consumed, which happens if the transition with rate k_1 takes place. The average flux in (4.9) can be defined as

$$J_A \equiv \lim_{t\to\infty} \frac{\langle Z_A(t)\rangle}{t\Omega}, \tag{4.10}$$

where T is the time interval and the brackets denote an average over stochastic trajectories. This time interval is large enough compared to relaxation times so that the stationary regime is probed. The diffusion coefficient (per volume) associated with $Z_A(t)$ is defined as

$$D_A \equiv \lim_{t\to\infty} \frac{\langle Z_A(t)^2\rangle - \langle Z_A(t)\rangle^2}{2t\Omega}. \tag{4.11}$$

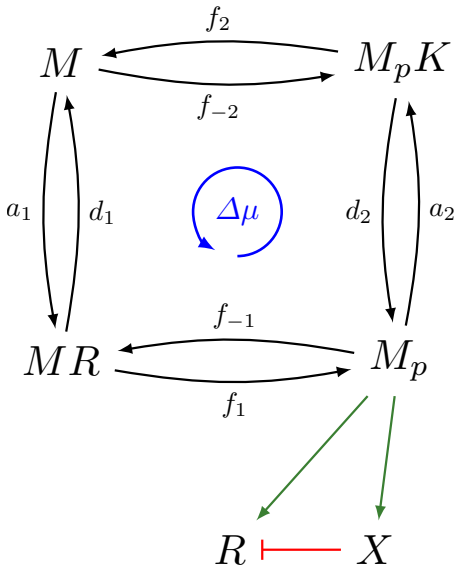

Figure 4.4: Activator-inhibitor model. The transition rates for the phosphorylation cycle of an enzyme M are represented with black arrows. The completion of the counter clock-wise cycle leads to the hydrolysis of one ATP, which liberates a free energy $\Delta\mu$. Green arrows represent activation and the red lines represent inhibition.

In Section 3.2.2, we have shown that the diffusion coefficient associated with the entropy production is

$$D_\sigma \equiv D_A \Delta\mu^2. \tag{4.12}$$

In Fig. 4.3(a) we show the diffusion coefficient D_σ as a function of $\Delta\mu$. It has a maximum close to the critical point that increases with the volume Ω. The maximum D_σ^c as a function of the volume Ω follows a power law with effective exponent 0.26 ± 0.01, as shown in Fig. 4.3(b). This finite-size scaling indicates that D_σ diverges at the critical point.

4.3 Activator-inhibitor model

4.3.1 Model definition

The activator-inhibitor model [122] is a more elaborate biochemical oscillator compared to the Brusselator. The model is depicted in Fig. 4.4. It consists of activators R, inhibitors X, enzymes M and phosphatases K interacting in a volume Ω. The external bath contains fixed concentrations of ATP, ADP, and P_i. The enzyme M can be in four different states (M, MR, M_pK, M_p), where M_p is the phosphorylated form of the enzyme.

For a phosphorylation reaction to take place, a R molecule must be bound to the enzyme M and for a dephosphorylation reaction to take place, a phosphatase K must be bound to the enzyme M. In the phosphorylation reaction, one ATP is transformed into an ADP, whereas in the dephosphorylation reaction a P_i is released in the solution. In the anti-clockwise cycle shown in Fig. 4.4, one ATP is transformed into ADP + P_i, hence the thermodynamic force associated with this cycle is

$$\Delta\mu = \mu_{\mathrm{ATP}} - \mu_{\mathrm{ADP}} - \mu_{\mathrm{P_i}} \equiv \ln \frac{a_1 a_2 f_1 f_2}{d_1 d_2 f_{-1} f_{-2}}, \tag{4.13}$$

where the rates a_i, d_i, $f_{\pm i}$ are given in Fig. 4.4.

Activators R and inhibitors X are related to this phosphorylation cycle in a feedback loop. The system contains a fixed concentration of substrate S, which is consumed (produced) when a R or X molecule is produced (consumed). The phosphorylated form of the enzyme M_p catalyzes both the production of R with a rate k_0 and the production of X with a rate k_3 (positive feedback). Inhibitors X degrade R with a rate k_2 (negative feedback) and can be spontaneously degraded with a rate k_4. For thermodynamic consistency, we must include reverse rates, which are given by δ_i, where $i = 0, \ldots, 4$. These reactions correspond to the lower part of Fig. 4.4 and can be written as

$$\begin{aligned}
M_p + S &\underset{\delta_0}{\overset{k_0}{\rightleftharpoons}} M_p + R, \\
S &\underset{\delta_1}{\overset{k_1}{\rightleftharpoons}} R, \\
X + R + \mathrm{ATP} &\underset{\delta_2}{\overset{k_2}{\rightleftharpoons}} X + S + \mathrm{ADP} + \mathrm{P_i}, \\
M_p + S &\underset{\delta_3}{\overset{k_3}{\rightleftharpoons}} M_p + X, \\
X + \mathrm{ATP} &\underset{\delta_4}{\overset{k_4}{\rightleftharpoons}} S + \mathrm{ADP} + \mathrm{P_i},
\end{aligned} \tag{4.14}$$

where generalized detailed balance [6] requires

$$\begin{aligned}
\delta_1 &= k_1 \delta_0 / k_0, \\
\delta_2 &= \mathrm{e}^{-\Delta\mu} k_2 k_0 / \delta_0, \\
\delta_4 &= \mathrm{e}^{-\Delta\mu} k_4 k_3 / \delta_3.
\end{aligned} \tag{4.15}$$

In contrast to the model from [122], we have added ATP consumption in the chemical reactions Eq. (4.14) for thermodynamic consistency.

An additional feature of the activator-inhibitor model in relation to the Brusselator is the competition for a scarce number of phosphatases $n_{K_{\mathrm{tot}}}$. In order for oscillations to

set in, this number must be at some intermediate optimal value. If $n_{K_{\text{tot}}} > n_{M_{\text{tot}}}$, then the phosphorylation cycle shown in Fig. 4.4 for different enzymes does not synchronize, since there is always free phosphatase to bind to the enzyme, which is necessary for the dephosphorylation reaction. If $n_{K_{\text{tot}}}$ is too small then only a few enzymes can complete their cycle in a synchronized way. This competition for a scarce resource is a common feature in more realistic biochemical oscillators, such as the model for KaiC oscillations in the next section.

The chemical master equation and the respective deterministic rate equations for this model are shown in Appendix 4.A. The state of the system is determined by a vector of the numbers of molecules n_i, with $i = R, X, K, M, MR, M_p, M_pK$. This vector is subjected to the constraints $n_{M_{\text{tot}}} = n_M + n_{MR} + n_{M_p} + n_{M_pK}$ and $n_{K_{\text{tot}}} = n_K + n_{M_pK}$, where $n_{M_{tot}}$ is the total number of enzymes and $n_{K_{tot}}$ is the total number of phosphatases. The volume of the system is Ω and concentrations are denoted by $[i] \equiv n_i/\Omega$. The concentration of enzymes is set to $[M_{tot}] = 10$, the concentration of phosphatases is $[K_{tot}] = 0.8$ and the concentration of substrates is $[S] = 1$. The rates are set to $k_0 = k_2 = k_3 = 0.02, k_1 = 0.008, k_4 = 0.01, \delta_0 = \delta_3 = 0.001, f_1 = f_2 = d_1 = d_2 = 0.3, a_1 = a_2 = 2$. The rates $\delta_1, \delta_2, \delta_4$ are computed with the generalized detailed balance relation in Eq. (4.13) and $f_{-1} = f_{-2} = 2e^{-\Delta\mu/2}$, where $\Delta\mu$ is the control parameter.

The chemical species that we observe in our numerical simulations is X, which, depending on $\Delta\mu$, can display biochemical oscillations. The fluctuating thermodynamic time-integrated current $Z_A(t)$ in this model is the total number of ATP consumed: if the reaction with rate f_1 (f_{-1}) in Fig. 4.4 takes place, the $Z_A(t)$ increases (decreases) by one. In addition, if the reactions with rate k_2, k_4 (δ_2, δ_4) in Eq. (4.14) take place, $Z_A(t)$ also increases (decreases) by one. The mean entropy production rate σ and its diffusion coefficient per volume D_σ are defined as in Eqs. (4.8) and (4.12), respectively.

4.3.2 Results

As shown in Fig. 4.5, the critical behavior of the activator-inhibitor model is qualitatively similar to the Brusselator. The first derivative of the mean entropy producion rate $\partial\sigma/\partial\Delta\mu$ has a discontinuity at the critical point in the thermodynamic limit, as obtained from the deterministic rate equations shown in Appendix 4.A. Furthermore, the stationary distribution $P(N_X)$ becomes bimodal above the critical point, which depends on the system size. The diffusion coefficient D_σ diverges at the critical point, as shown in Fig. 4.6. The effective exponent related to the finite-size scaling of the maximum diffusion coefficient D_σ^c is 0.41 ± 0.03, which is different from the one found in the Brusselator.

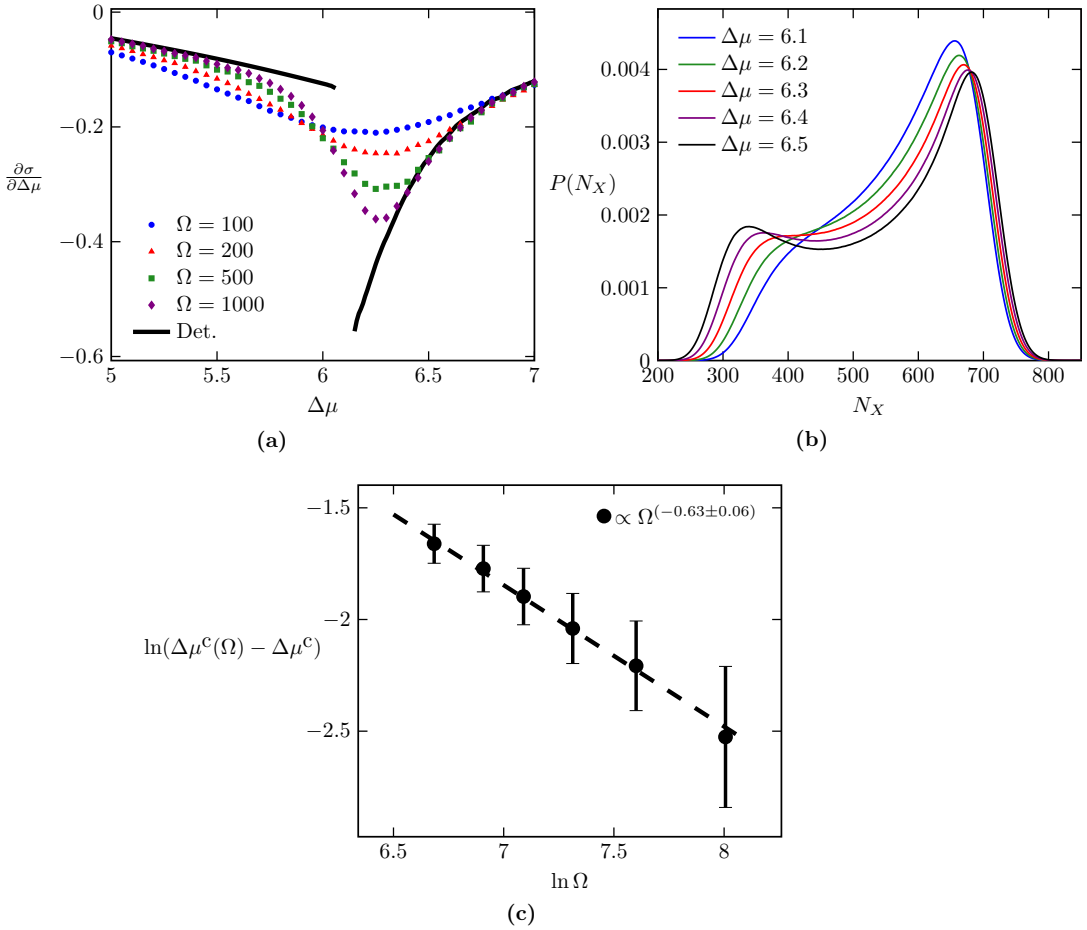

Figure 4.5: Phase transition for the activator-inhibitor model. (a) First derivative of the mean entropy production rate $\partial\sigma/\partial\Delta\mu$ as a function of $\Delta\mu$. The critical point is $\Delta\mu^c \simeq 6.1$. (b) Stationary probability distribution of N_X for different values of $\Delta\mu$ and $\Omega = 100$. The distribution becomes bimodal above the critical point, which is $\Delta\mu \simeq 6.3$ for this finite Ω. (c) Difference between the point at which the distribution in (b) becomes bimodal in a finite system $\Delta\mu^c(\Omega)$ and the critical point $\Delta\mu^c \simeq 6.1$ obtained with the deterministic rate equations.

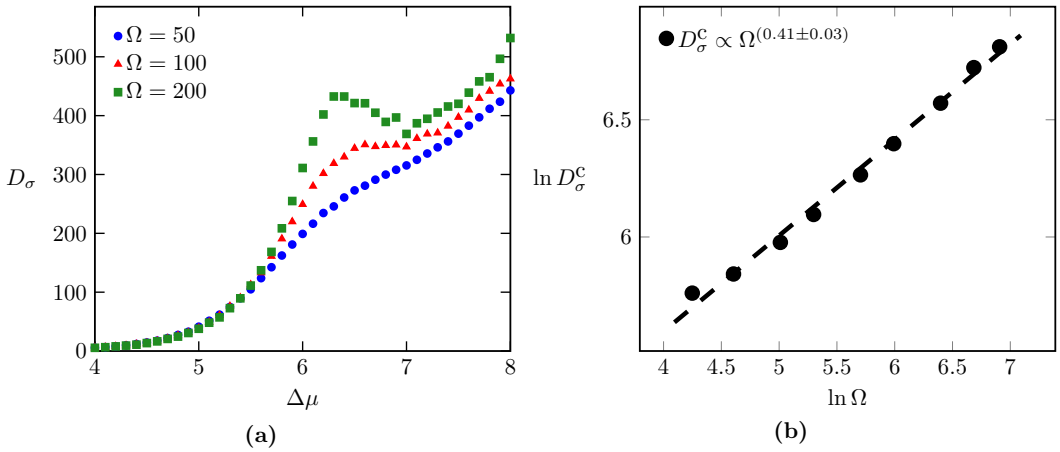

Figure 4.6: Diffusion coefficient for the activator-inhibitor model. (a) Diffusion coefficient D_σ as a function of $\Delta\mu$. (b) Maximum diffusion coefficient D_σ^c as a function of Ω.

4.4 KaiC model

4.4.1 Model definition

There are several models for the oscillations of the phosphorylation level of KaiC proteins (see [155] for a summary). Here, we analyze a modified version of a model introduced in [150]. In particular, we make all transitions reversible for thermodynamic consistency.

The model contains KaiC molecules and KaiA molecules in a volume Ω. Each KaiC molecule can be in 14 different states denoted by C_j and $\widetilde{C}_j$, where $j = 0, 1, \ldots, 6$. The variables j indicates the phosphorylation level of the molecule, which has 6 phosphorylation sites. The state C_j indicates the molecule is active and the state $\widetilde{C}_j$ indicates the molecule is inactive. The free energy of state C_j minus the free energy of state $\widetilde{C}_j$ is given by

$$\Delta E_j = -E + jE/3. \tag{4.16}$$

If the molecule is active then a phosphorylation reaction can happen and if the molecule is inactive a dephosphorylation reaction can happen.

An essential feature of the model is that a KaiA molecule must bind to the KaiC molecule for a phosphorylation reaction. The KaiA molecules play a role similar to the phosphatases in the activator-inhibitor model, i.e., it is a scarce resource that synchronizes the phosphorylation cycle of different KaiC molecules. The dissociation constant for the binding of an A to a KaiC molecule in state C_j is $K_j = K_0 a^j$. The constant $a > 1$ makes the dissociation constant an increasing function of j, which is necessary for the onset of biochemical oscillations [150].

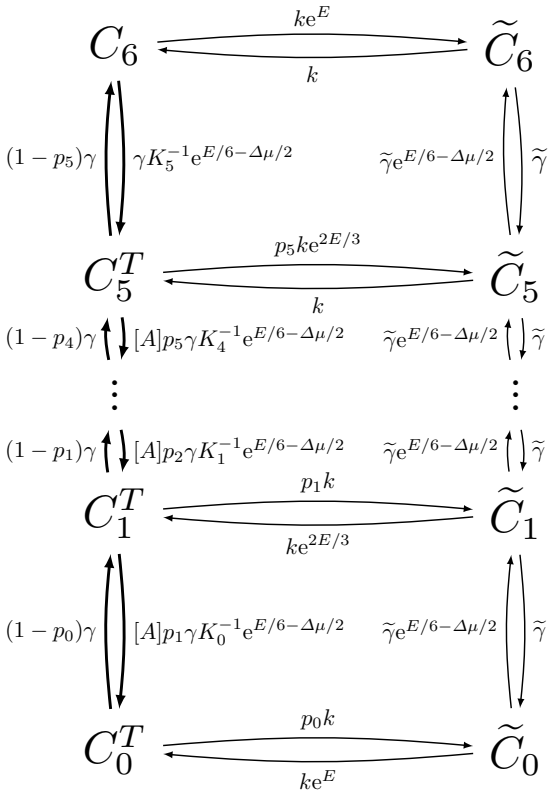

Figure 4.7: KaiC model with differential affinity. $C_j^T = C_j + AC_j$ are the active KaiC states with j bound phosphates, which can be bound to a KaiA(A). We assume fast binding and unbinding of KaiA with C_j, which is characterized by a dissociation constant $K_j = K_0 a^j$ and the fraction of unbound states $p_j = K_j/(K_j + [A])$. Note that the fully phosphorylated state C_6 cannot bind to A. $\widetilde{C}_j$ are the inactive states with j bound phosphates. The transition rate from C_j to $\widetilde{C}_j$ is $ke^{\chi E(j-3)/3}$ and the transition rate from $\widetilde{C}_j$ to C_j is $ke^{\tilde{\chi}E(3-j)/3}$, where χ ($\tilde{\chi}$) is an indicator function that is 0 (1) for $j = 0, 1, 2, 3$ and 1(0) for $j = 4, 5, 6$. See Appendix 4.B for the chemical master equation and rate equations.

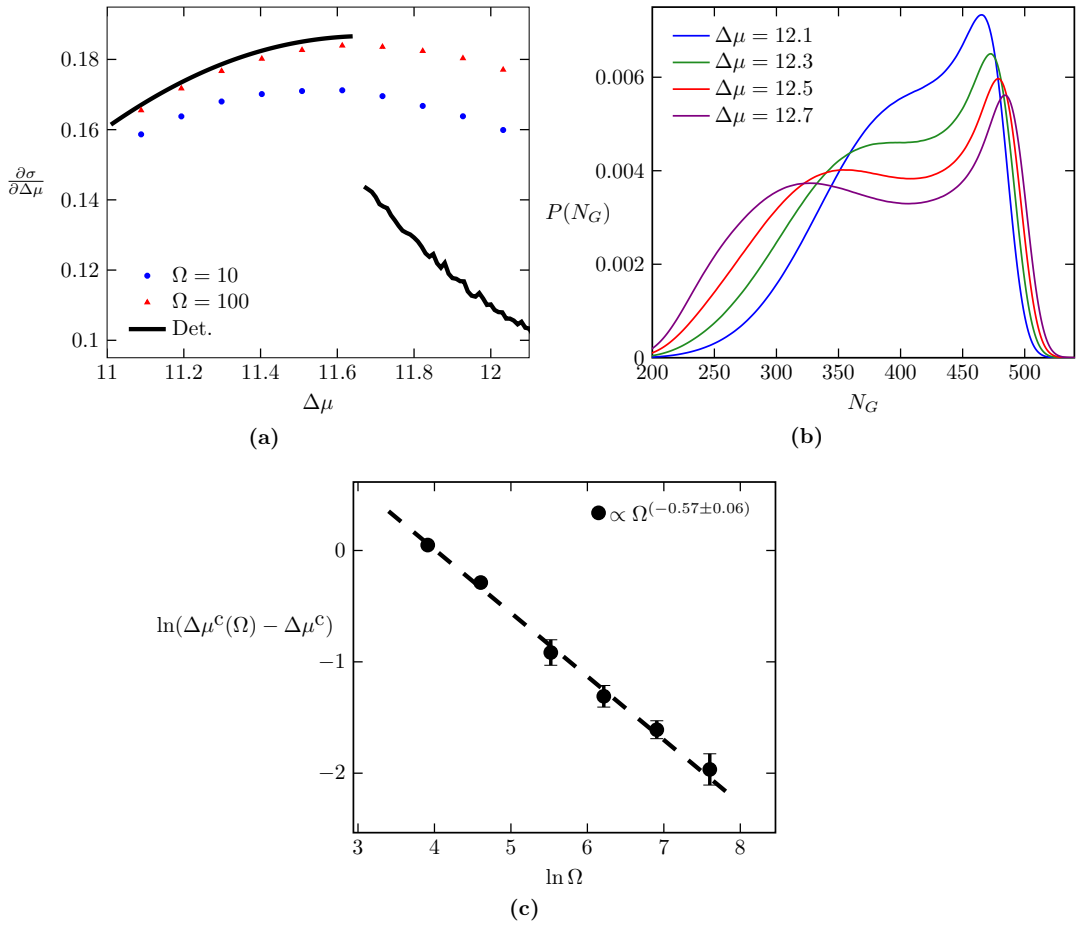

Figure 4.8: Phase transition for the KaiC model. (a) First derivative of the mean entropy production rate $\partial\sigma/\partial\Delta\mu$ as a function of $\Delta\mu$. The critical point is $\Delta\mu^c \simeq 11.6$. (b) Stationary probability distribution of the phosphorylation level n_G for different values of $\Delta\mu$ and $\Omega = 100$. The distribution becomes bimodal above the critical point, which is $\Delta\mu \simeq 12.3$ for this finite Ω. (c) Difference between the point at which the distribution in (b) becomes bimodal in a finite system $\Delta\mu^c(\Omega)$ and the critical point $\Delta\mu^c \simeq 11.6$ obtained with the deterministic rate equations.

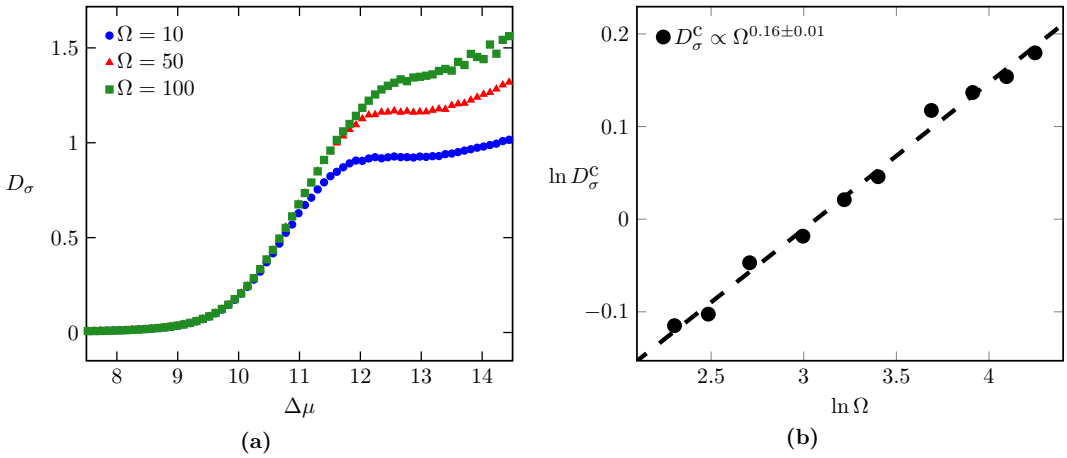

Figure 4.9: Diffusion coefficient for the KaiC model. (a) Diffusion coefficient D_σ as a function of $\Delta\mu$. (b) Maximum diffusion coefficient D_σ^c as a function of Ω.

The transition rates for the model are as follows. KaiA(A) can bind with rate α to active states which are partially phosphorylated ($j = 0, ..., 5$) to form a complex AC_j. The transition rate for the unbinding of A from a KaiC in state AC_j is αK_j. The rate for a phosphorylation reaction from AC_j to C_{j+1} is γ, while the rate for the reverse reaction is rate $\gamma K_j^{-1} e^{E/6-\Delta\mu/2}$. Dephosphorylation from $\widetilde{C}_{j+1}$ to $\widetilde{C}_j$ occurs with rate $\widetilde{\gamma}$, whereas the rate for the reversed reaction is $\widetilde{\gamma} e^{E/6-\Delta\mu/2}$. The transition rate from C_j to $\widetilde{C}_j$ is $k e^{\chi E(j-3)/3}$ and the transition rate from $\widetilde{C}_j$ to C_j is $k e^{(1-\chi)E(3-j)/3}$, where χ_j is an indicator function that is 0 for $j = 0, 1, 2, 3$ and 1 for $j = 4, 5, 6$. This transition rates are thermodynamically consistent since the affinity of a phosporylation cycle is $\Delta\mu$. The chemical master equation for this model is shown in Appendix 4.B.

We assume fast binding and unbinding kinetics of KaiA and introduce coarse-grained states $C_j^T = C_j + AC_j$ for $j = 0, .., 5$. The transition rates for this coarse-grained model are shown in Fig. 4.7. Within this assumption of timescale separation, the entropy production associated with this coarse-grained model remains the same as the entropy production of the full model [65]. The fraction of unbound states among C_j^T is $p_j = K_j/(K_j + [A])$, where [A] is the concentration of free KaiA. This concentration fulfills the constraint

$$[A_{\text{tot}}] = [A] + \sum_{j=0}^{6} \frac{[A]\left[C_j^T\right]}{K_j + [A]}, \tag{4.17}$$

where $[A_{\text{tot}}]\,\Omega$ is the total number of KaiA molecules. The total number of KaiC is $[C_{\text{tot}}]\Omega$, which leads to the constraint

$$[C_{\text{tot}}] = \sum_{j=0}^{6} \left([C_j] + [\widetilde{C}_j]\right). \tag{4.18}$$

The chemical master equation and deterministic rate equations for this coarse-grained model are also shown in Appendix 4.B.

The concentration of KaiC is set to $[C_{\text{tot}}] = 1$, while the concentration of KaiA is set to $[A_{\text{tot}}] = 0.04$. The parameters determining the transition rates are set to $\gamma = 10, \widetilde{\gamma} = 1, E = 20, k = e^{-E}, a = 10, K_0 = (1/3)10^{-7}$. The thermodynamic force $\Delta\mu$ is kept as a free parameter. In our simulations, we monitor the phosphorylation level of the KaiC system, which is defined as

$$n_G = \sum_{j=1}^{6} j \left([C_j] + [\widetilde{C}_j] \right). \tag{4.19}$$

Depending on $\Delta\mu$ this quantity can exhibit oscillations. The probability distribution $P(n_G)$ is the steady state probability of this observable. The fluctuating thermodynamic time-integrated current $Z_A(t)$ in this model is the total of number of ATP consumed. For the KaiC molecule in the active state, if a transition that increases j takes place $Z_A(t)$ increases by one and if the reversed transition takes place $Z_A(t)$ decreases by one. The mean entropy production rate σ and its diffusion coefficient per volume D_σ are defined as in Eqs. (4.8) and (4.12), respectively.

4.4.2 Results

The critical behavior of the present model is similar to the critical behavior observed with the Brusselator and the activator-inhibitor model, as shown in Fig. 4.8. The first derivative of the mean entropy production rate $\partial\sigma/\partial\Delta\mu$ has a discontinuity at the critical point in the thermodynamic limit and the stationary distribution $P(n_G)$ becomes bimodal above the critical point. For this model, concerning the quantity $\partial\sigma/\partial\Delta\mu$, we could not numerically access systems sizes that are large enough for a better agreement between finite systems and the result from deterministic rate equations in Fig. 4.8(a). The diffusion coefficient D_σ diverges at the critical point, as shown in Fig. 4.9. The effective exponent related to the finite-size scaling of the maximum D_σc as a function of the the volume Ω is 0.16 ± 0.01, which is smaller then the effective exponents found for the other two models.

4.5 Criteria for the precision of biochemical oscillations

Biochemical oscillations can occur in finite systems that display large fluctuations. In this section, we compare two quantities that can quantify the precision of biochemical oscillations, the number of coherent oscillations and the Fano factor associated with the

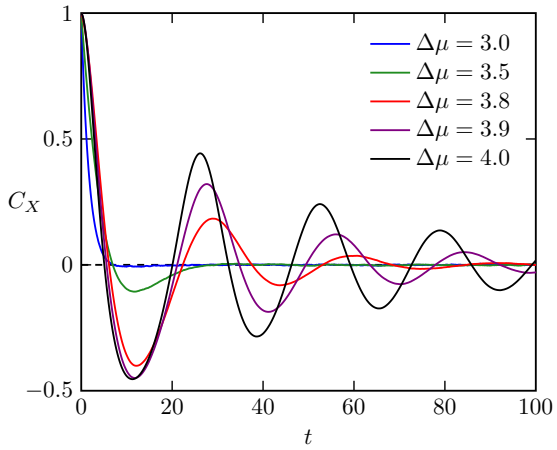

Figure 4.10: Correlation function for species X (Eq. (4.20)) in the Brusselator for different $\Delta\mu$. The volume is $\Omega = 10^3$.

entropy current. The first quantity is related to the correlation function

$$C_X(t) \equiv \frac{\left\langle \left(N_\mathrm{X}(t) - \langle N_\mathrm{X}\rangle\right)\left(N_\mathrm{X}(0) - \langle N_\mathrm{X}\rangle\right)\right\rangle}{\langle N_\mathrm{X}^2\rangle - \langle N_\mathrm{X}\rangle^2}. \tag{4.20}$$

where $N_X(t)$ is the number of X molecules at time t, the average number of molecules is $\langle N_X\rangle \equiv \sum_{N_X} N_X P(N_X)$. Note that for the KaiC model the same expression is valid with n_G instead of N_X. In our simulations of all three models, the system starts at the initial time $t = 0$ in the stationary state. In Fig. 4.10, we show the correlation function for species X in the Brusselator. Above the critical point, it displays oscillations with an exponentially decreasing amplitude . The number of coherent oscillations $\mathcal{N}$ is defined as the decay time divided by the period. It gives the typical number of oscillations for which two different stochastic realizations would remain coherent with each other. Therefore, a non-zero $\mathcal{N}$ is a signature of biochemical oscillations, since it is a result of a non-zero imaginary part of the first excited eigenvalue of the stochastic matrix [123, 144].

As shown in Fig. 4.11, for a finite system, $\mathcal{N}$ becomes non-zero above some critical value of $\Delta\mu$. It is difficult to evaluate $\mathcal{N}$ numerically close to the critical point. If the number of coherent oscillations is too small, it is not possible to obtain the period of oscillation from direct numerical evaluation of the correlation function. Hence, with our numerical results, we cannot determine whether $\mathcal{N}$ has a jump or approaches zero smoothly at the critical point. However, for simple three-state models it is possible to evaluate $\mathcal{N}$ by calculating the first non-zero eigenvalue of the stochastic matrix [123, 144]. In this analytical case, $\mathcal{N}$ approaches zero smoothly, without a discontinuity. Within our numerical simulations, the chemical force for which $\mathcal{N}$ becomes non-zero is smaller then the critical $\Delta\mu$ for which the probability distribution of the oscillating species becomes bimodal, for all three models.

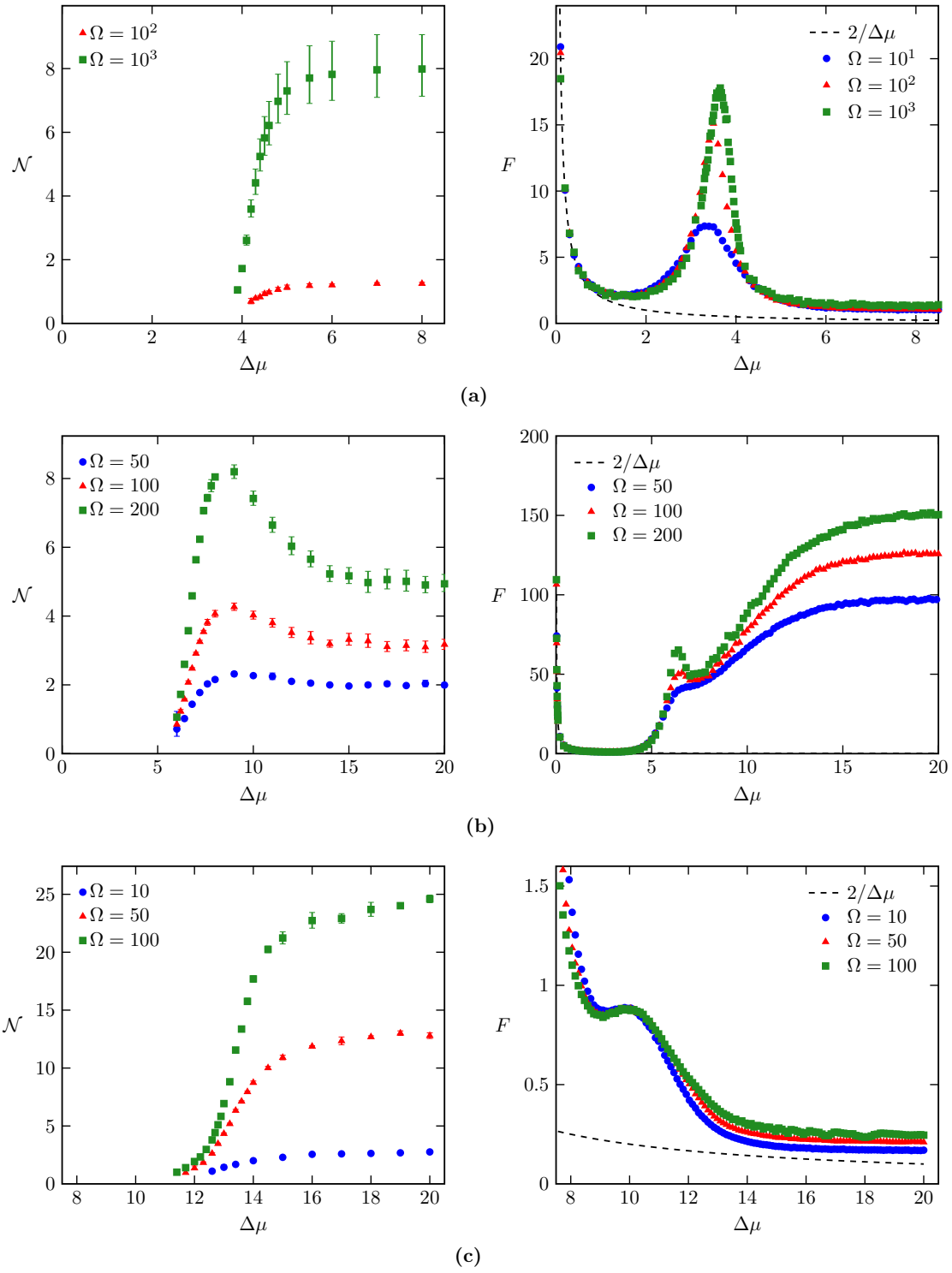

Figure 4.11: Comparison between the number of coherent oscillations $\mathcal{N}$ and the Fano factor F_A. (a) Brusselator ($\mathcal{N}$ could not be measured for $\Omega = 10$). (b) Activator-inhibitor model. (c) Model for KaiC oscillations. The right panels also show the lower bound (4.22) for F_A.

For the Brusselator, the scaling of $\mathcal{N}$ with the volume Ω has been analyzed for both $\Delta\mu$ close to the critical point [146, 147] and $\Delta\mu$ above the critical point [145]. In the first case, the scaling $\mathcal{N} \propto \Omega^{1/2}$ has been observed, while in the second $\mathcal{N} \propto \Omega^{1}$ has been obtained with the linear noise approximation. We have observed that close to the critical point our results for all three models, which are shown in Fig. 4.11, agree with the scaling $\mathcal{N} \propto \Omega^{1/2}$. Above the critical point, we find an exponent smaller than 1, which suggest that the linear noise approximation, which is valid for large Ω, does not apply to the system sizes we have considered. Furthermore, in [147] Xiao *et al.* have shown that $\mathcal{N}$ becomes non-zero below the critical point of the Brusselator. We also observe a non-zero $\mathcal{N}$ below $\Delta\mu^{c}$ for the activator-inhibitor and KaiC models.

The Fano factor F_A associated with the thermodynamic current J_A is defined as

$$F_A \equiv \lim_{t\to\infty} \frac{\langle Z_A(t)^2\rangle - \langle Z_A(t)\rangle^2}{\langle Z_A(t)\rangle} = \frac{2D_A}{J_A}, \tag{4.21}$$

which is a dimensionless parameter that quantifies the fluctuations of the thermodynamic time-integrated current $Z_A(t)$. A small Fano factor means that the current is precise. In equilibrium, where there are no biochemical oscillations, the Fano factor diverges, since the average current is zero. Hence, this quantity is consistent with the absence of biochemical oscillations in equilibrium. Due to the divergence of the diffusion coefficient $D_\sigma = D_A\Delta\mu^2$, the Fano factor also diverges at the critical point in the thermodynamic limit.

The thermodynamic uncertainty relation[20] implies the bound

$$F_A \geq \frac{2}{\Delta\mu}. \tag{4.22}$$

The highest precision achievable increases with the thermodynamic force $\Delta\mu$. Therefore, a large $\Delta\mu$ is a necessary condition for a small Fano factor. This inequality is illustrated in Fig. 4.11. Close to the critical point, F_A is substantially larger than the bound with a local maximum that increases with the volume Ω.

Two main motivations to consider the Fano factor as a metric for the precision of biochemical oscillations are as follow. First, in light of the thermodynamic uncertainty relation, the Fano factor is a natural observable to relate precision with thermodynamics. Second, for a simple unicyclic mode, the Fano factor can be related to the cycle completion time (i.e. the period of oscillations) [124]. Furthermore, the Fano factor associated with the ATP consumption seems particularly appealing to quantifying the precision of oscillations in the KaiC model, for which we consider oscillations of the phosphorylation level.

However, our results in Fig. 4.11 indicate that F_A does not appropriately quantify the precision of biochemical oscillations. In particular, for the activator-inhibitor in Fig. 4.11(b), the Fano factor is minimal below the critical point where there are no

biochemical oscillations. In general, since the Fano factor diverges at the critical point in the thermodynamic limit, a system that is large enough should display a smaller F_A below the critical point as compared to the F_A is some region above the critical point. Note that it is more difficult to observe this effect in the KaiC model, due to the smaller effective exponent associated with the divergence of D_σ. For the Brusselator and KaiC model, above the crossover to oscillatory behavior, an increase in $\Delta\mu$ leads to an increase in $\mathcal{N}$ and a decrease in F_A. However, for the activator-inhibitor model, there is a region around $\Delta\mu \simeq 8$ where both $\mathcal{N}$ and F_A increase with $\Delta\mu$. Hence, for this last model, even if we restrict to a region where biochemical oscillations exist, F_A does not appropriately correlate with the precision quantified by $\mathcal{N}$.

4.6 Conclusion

We have characterized a generic phase transition in biochemical oscillators. A control parameter for this phase transition is the thermodynamic force that drives the system out of equilibrium. The stationary distribution of the oscillating chemical species becomes bimodal above the critical force, the first derivative of mean entropy production rate with respect to the force is discontinuous at criticality, its diffusion coefficient diverges at criticality, and the number of coherent oscillations, which is a signature of biochemical oscillations, becomes nonzero above the critical point. We have estimated effective exponents for the divergence of the diffusion coefficient with a limited range of volumes, a more reliable calculation of this exponent remains an open problem.

Three different models that display a limit cycle in the deterministic limit have been investigated. We expect that models for biochemical oscillators with this feature display qualitatively the same critical behavior. For models with purely stochastic biochemical oscillations [156], i.e., models that do not have a limit cycle in the deterministic limit, this phase transition remains unexplored so far. The critical behavior of the entropy production has also been investigated in several different models with nonequilibrium phase transitions [91, 92, 94–97, 157]. The present case constitutes an example for which the first derivative of entropy production, which is the flux multiplied by the force, displays a discontinuity at the phase transition. We expect that the diffusion coefficient associated with the fluctuating entropy production diverges at the critical point for other models with nonequilibrium phase transitions.

The precision of biochemical oscillations can be characterized by the Fano factor F_A associated with the thermodynamic flux (or related first passage variables [124]) and the number of coherent oscillations $\mathcal{N}$. We have shown that the Fano factor F_A can indicate high precision even if there are no biochemical oscillations. This result suggests that F_A does not quantify the precision of biochemical oscillations in a reliable way. We anticipate that this problem will also happen with probability currents that are different from the thermodynamic flux, which could be considered as more suitable for quantifying

the precision of biochemical oscillations, since their diffusion coefficient will also diverge at criticality. It remains to be seen whether and how the phase transition that we have characterized here can be used to understand the operation and design principles of biochemical oscillators.

Appendices to chapter 4

4.A Activator-inhibitor model

The time evolution of $P(N_R, N_X, N_K, N_{\mathbf{M}}, t)$, the probability to find the system in state $(N_R, N_X, N_K, N_{\mathbf{M}})$ at time t, is governed by the chemical master equation

$$\begin{aligned}\frac{\partial}{\partial t}P(N_R, N_X, N_K, N_{\mathbf{M}}, t) = \Big\{ & k_0[S](\mathcal{E}_R^- - 1)N_{M_p} + \Omega k_1[S](\mathcal{E}_R^- - 1) + k_3[S](\mathcal{E}_X^- - 1)N_{M_p} \\ & + k_2/\Omega\left(\mathcal{E}_R^+ - 1\right) N_X N_R + k_4\left(\mathcal{E}_X^+ - 1\right) N_X \\ & + \delta_0/\Omega\left(\mathcal{E}_R^+ - 1\right) N_{M_p} N_R + \delta_1\left(\mathcal{E}_R^+ - 1\right) N_R \\ & + \delta_2 N_X[S](\mathcal{E}_R^- - 1) + \delta_3/\Omega\left(\mathcal{E}_X^+ - 1\right) N_{M_p} N_X + \Omega\delta_4[S](\mathcal{E}_X^- - 1) \\ & + a_1/\Omega\left(\mathcal{E}_M^+\mathcal{E}_{MR}^- - 1\right)(N_R - N_{MR})N_M \\ & + a_2/\Omega\left(\mathcal{E}_K^+\mathcal{E}_{M_p}^+\mathcal{E}_{M_pK}^- - 1\right) N_{M_p} N_K \\ & + d_1\left(\mathcal{E}_M^-\mathcal{E}_{MR}^+ - 1\right) N_{MR} \\ & + d_2\left(\mathcal{E}_K^-\mathcal{E}_{M_p}^-\mathcal{E}_{M_pK}^+ - 1\right) N_{M_pK} \\ & + f_1\left(\mathcal{E}_{MR}^+\mathcal{E}_{M_p}^- - 1\right) N_{MR} \\ & + f_2\left(\mathcal{E}_K^-\mathcal{E}_M^-\mathcal{E}_{M_pK}^+ - 1\right) N_{M_pK} \\ & + f_{-1}/\Omega\left(\mathcal{E}_{MR}^-\mathcal{E}_{M_p}^+ - 1\right)(N_R - N_{MR})N_{M_p} \\ & + f_{-2}/\Omega\left(\mathcal{E}_K^+\mathcal{E}_M^+\mathcal{E}_{M_pK}^- - 1\right) N_K N_M \\ & \Big\} P(N_R, N_X, N_K, N_{\mathbf{M}}, t),\end{aligned} \tag{4.23}$$

where we define step operators as $\mathcal{E}_i^{\pm}P(N_R, N_X, N_K, N_{\mathbf{M}}, t) \equiv P(..., N_i \pm 1, ..., t)$ and $\mathcal{E}_i^{\pm}N_i \equiv (N_i \pm 1)$ for $i = R, X, K, M, MR, M_p, M_pK$.

From the master Eq. (4.23), we obtain the equations for the time evolution of the concentrations in the deterministic limit [122], which reads

$$
\begin{aligned}
\frac{d[R]}{dt} &= k_0[M_p][S] + k_1[S] - k_2[X][R] - \delta_0[M_p][R] - \delta_1[R] + \delta_2[X][S] \\
\frac{d[X]}{dt} &= k_3[M_p][S] - k_4[X] - \delta_3[M_p][X] + \delta_4[S] \\
\frac{d[M]}{dt} &= f_2[M_pK] + d_1[MR] - a_1[M]([R] - [MR]) - f_{-2}[M][K] \\
\frac{d[MR]}{dt} &= a_1[M]([R] - [MR]) + f_{-1}[M_p]([R] - [MR]) - (f_1 + d_1)[MR] \\
\frac{d[M_p]}{dt} &= f_1[MR] + d_2[M_pK] - a_2[M_p][K] - f_{-1}[M_p]([R] - [MR]) \\
\frac{d[M_pK]}{dt} &= a_2[M_p][K] + f_{-2}[M][K] - f_2[M_pK] - d_2[M_pK] \\
[K_{\text{tot}}] &= [K] + [M_pK].
\end{aligned}
\tag{4.24}
$$

Note that $[R]$ is the total concentration of activators and $([R]-[MR])$ is the concentration of free activators which can bind an enzyme M.

4.B KaiC model

For the KaiC model, the state of the system is determined by the number of free active KaiC $N_{\mathbf{C}} = (N_{C_0}, N_{C_1}, ..., N_{C_6})$, inactive KaiC $N_{\widetilde{\mathbf{C}}} = (N_{\widetilde{C}_0}, N_{\widetilde{C}_1}, ..., N_{\widetilde{C}_6})$, active KaiC bound with a KaiA $N_{\mathbf{AC}} = (N_{AC_0}, N_{AC_1}, ..., N_{AC_5})$ and free KaiA N_A. The time evolution of $P(N_{\mathbf{C}}, N_{\widetilde{\mathbf{C}}}, N_{\mathbf{AC}}, N_A, t)$, the probability to find the system in state $(N_{\mathbf{C}}, N_{\widetilde{\mathbf{C}}}, N_{\mathbf{AC}}, N_A, t)$ at time t, is governed by the chemical master equation

$$
\begin{aligned}
\frac{\partial}{\partial t}P(N_{\mathbf{C}}, N_{\widetilde{\mathbf{C}}}, N_{\mathbf{AC}}, N_A, t) = \Big\{ & \sum_{j=0}^{5} \gamma \left(\mathcal{E}^{+}_{AC_j} \mathcal{E}^{-}_{C_{j+1}} - 1 \right) N_{AC_j} \\
& + \sum_{j=0}^{5} \gamma K_j^{-1} e^{E/6 - \Delta\mu/2} \left(\mathcal{E}^{+}_{C_{j+1}} \mathcal{E}^{-}_{AC_j} - 1 \right) N_{C_{j+1}} \\
& + \sum_{j=0}^{5} \widetilde{\gamma} e^{E/6 - \Delta\mu/2} \left(\mathcal{E}^{+}_{\widetilde{C}_j} \mathcal{E}^{-}_{\widetilde{C}_{j+1}} - 1 \right) N_{\widetilde{C}_j} \\
& + \sum_{j=0}^{5} \widetilde{\gamma} \left(\mathcal{E}^{+}_{\widetilde{C}_{j+1}} \mathcal{E}^{-}_{\widetilde{C}_j} - 1 \right) N_{\widetilde{C}_{j+1}} \\
& + \sum_{j=0}^{5} \alpha/\Omega \left(\mathcal{E}^{+}_{C_j} \mathcal{E}^{+}_{A} \mathcal{E}^{-}_{AC_j} - 1 \right) N_{C_j} N_A \\
& + \sum_{j=0}^{5} \alpha K_j \left(\mathcal{E}^{+}_{AC_j} \mathcal{E}^{-}_{C_j} \mathcal{E}^{-}_{A} - 1 \right) N_{AC_j} \\
& + \sum_{j=0}^{6} k e^{\chi E(3-j)/3} \left(\mathcal{E}^{+}_{C_j} \mathcal{E}^{-}_{\widetilde{C}_j} - 1 \right) N_{C_j} \\
& + \sum_{j=0}^{6} k e^{\widetilde{\chi} E(j-3)/3} \left(\mathcal{E}^{+}_{\widetilde{C}_j} \mathcal{E}^{-}_{C_j} - 1 \right) N_{\widetilde{C}_j} \\
& \Big\} P(N_{\mathbf{C}}, N_{\widetilde{\mathbf{C}}}, N_{\mathbf{AC}}, N_A, t),
\end{aligned}
\tag{4.25}
$$

where the step operators are defined as as $\mathcal{E}^{\pm}_{j} P(N_{\mathbf{C}}, N_{\widetilde{\mathbf{C}}}, N_{\mathbf{AC}}, N_A, t) \equiv P(..., N_j \pm 1, ..., t)$ and $\mathcal{E}^{\pm}_{j} N_j \equiv \left(N_j \pm 1\right)$ for $j = C_0, ..., C_6, \widetilde{C}_0, ..., \widetilde{C}_6, AC_0, ..., AC_5, A$.

For the fast binding and unbinding kinetics of KaiA, we introduce coarse-grained active states $N_{C_j^T} = N_{C_j} + N_{AC_j}$, where $j = 0, 5$ and $N_{C_6^T} = N_{C_6}$ (KaiA does not bind to C_6). We define the unbound fraction of active states as $p_j = K_j/(K_j + [A])$. The concentration of free KaiA $[A]$ is given implicitly by the constraint

$$
[A] + \sum_{j=0}^{5} \frac{[A] \left[C_j^T \right]}{K_j + [A]} = [A_{\mathrm{tot}}] \,. \tag{4.26}
$$

The state of the coarse-grained system is determined by the total number of active KaiC $N_{\mathbf{C}^T} = (N_{C_1^T}, ..., N_{C_6^T})$ and inactive KaiC $N_{\widetilde{\mathbf{C}}} = (N_{\widetilde{C}_1}, ..., N_{\widetilde{C}_6})$. The time evolution of

$P(N_{\mathbf{C}^T}, N_{\widetilde{\mathbf{C}}}, t)$, the probability to find the system in state $(N_{\mathbf{C}^T}, N_{\widetilde{\mathbf{C}}}, t)$ at time t, is governed by the following chemical master equation

$$\begin{aligned}\frac{\partial}{\partial t}P(N_{\mathbf{C}}, N_{\widetilde{\mathbf{C}}}, t) = \Bigg\{ & \sum_{j=0}^{5}(1-p_j)\gamma\left(\mathcal{E}^{+}_{N_{C_j^T}}\mathcal{E}^{-}_{N_{C_{j+1}^T}} - 1\right)N_{C_j^T} \\ & + \sum_{j=0}^{4}[A]p_j\gamma K_j^{-1}e^{E/6-\Delta\mu/2}\left(\mathcal{E}^{+}_{N_{C_{j+1}^T}}\mathcal{E}^{-}_{N_{C_j^T}} - 1\right)N_{C_{j+1}^T} \\ & + \gamma K_5^{-1}e^{E/6-\Delta\mu/2}\left(\mathcal{E}^{+}_{N_{C_6^T}}\mathcal{E}^{-}_{N_{C_5^T}} - 1\right)N_{C_6^T} \\ & + \sum_{j=0}^{5}\widetilde{\gamma}e^{E/6-\Delta\mu/2}\left(\mathcal{E}^{+}_{\widetilde{C}_j}\mathcal{E}^{-}_{\widetilde{C}_{j+1}} - 1\right)N_{\widetilde{C}_j} \\ & + \widetilde{\gamma}\left(\mathcal{E}^{+}_{\widetilde{C}_{j+1}}\mathcal{E}^{-}_{\widetilde{C}_j} - 1\right)N_{\widetilde{C}_{j+1}} \\ & + \sum_{j=0}^{6}p_j k e^{\chi E(3-j)/3}\left(\mathcal{E}^{+}_{C_j}\mathcal{E}^{-}_{\widetilde{C}_j} - 1\right)N_{C_j} \\ & + k e^{\widetilde{\chi}E(j-3)/3}\left(\mathcal{E}^{+}_{\widetilde{C}_j}\mathcal{E}^{-}_{C_j} - 1\right)N_{\widetilde{C}_j} \\ & \Bigg\}P(N_{\mathbf{C}}, N_{\widetilde{\mathbf{C}}}, t).\end{aligned} \tag{4.27}$$

From the coarse-grained master Eq. (4.27), we obtain the equations for the time evolution of the concentrations in the deterministic limit

$$\begin{aligned}\frac{d\left[C_j^T\right]}{dt} = & (1-p_{j-1})\gamma\left[C_{j-1}^T\right] + p_{j+1}\gamma K_{j+1}^{-1}\mathrm{e}^{E/6-\Delta\mu/2}\,[A]\left[C_{j+1}^T\right] \\ & - (1-p_j)\gamma\left[C_j^T\right](1-\delta_{j,6}) - p_j\gamma K_j^{-1}\mathrm{e}^{E/6-\Delta\mu/2}\,[A]\left[C_j^T\right](1-\delta_{j,0}) \\ & + k\mathrm{e}^{\widetilde{\chi}E(3-j)/3}\left[\widetilde{C}_j\right] - p_j k\mathrm{e}^{\chi E(j-3)/3}\left[C_j^T\right] \\ \frac{d\left[\widetilde{C}_j\right]}{dt} = & \widetilde{\gamma}\left(\left[\widetilde{C}_{j+1}\right] - \left[\widetilde{C}_j\right](1-\delta_{j,0})\right) - \widetilde{\gamma}\mathrm{e}^{\Delta\mu/2}\left(\left[\widetilde{C}_{j-1}\right] - \left[\widetilde{C}_j\right](1-\delta_{j,6})\right) \\ & + p_j k\mathrm{e}^{\chi E(j-3)/3}\left[C_j\right] - k\mathrm{e}^{\widetilde{\chi}E(3-j)/3}\left[\widetilde{C}_j\right],\end{aligned} \tag{4.28}$$

where $[A]$ is given by Eq. (4.26).

5 Concluding perspective

Thermodynamics is a universal theory for describing exchange processes. All energy exchanges such has work, heat or the exchange of matter must obey constraints imposed by the laws of thermodynamics that hold since the 19th century. This formalism has substantially improved to its latest formulation, stochastic thermodynamics, which can deal with small systems that are subject to large external noise and that operate far from equilibrium. In Chapter 2, we have presented a modern formulation of stochastic thermodynamics and have introduced chemical reaction networks.

In the main chapters of this thesis, we have characterized the behavior of the entropy current and its fluctuations at first- and second-order phase transition in homogenous chemical reaction networks. In the former case, we have shown that the mean rate of entropy production has a discontinuity at the bistable point and that the fluctuations have an exponential volume-dependence. In the latter case, we have shown that the first derivative of the mean rate of entropy production is discontinuous with respect to the thermodynamic force at the critical point and that fluctuations diverge as a power-law with the volume.

In Chapter 3 we have investigated first-order phase transitions with a paradigmatic model for biochemical switches, the Schlögl model. We find that the diffusion coefficient associated with entropy current diverges as an exponential function of the volume. This particular behavior differs from the fluctuations of the number of chemical species, which diverges as a power-law at the phase transition [158]. We have performed a coarse-graining of the chemical master equation and have obtained an effective two-state model. The exponential volume-dependence is due to a difference in the rate of entropy production at the fixed-points and the coarse-grained transition rates, which depend exponentially on the system size. The derivation of these transition rates starting with the chemical master equation can be challenging. As an alternative approach, we could solve the Fokker-Planck equation and obtain an analytical expression for the diffusion coefficient valid for large system sizes. At the bistable point, the exponential prefactor is given by the height of the effective potential barrier separating the two fixed-points.

We have considered a simple chemical network with a single degree of freedom. Systems with multiple chemical species, such as the interlinked GTPase cascade [45], could also be investigated using our framework. With a suitable change of reaction coordinates, one can always make the system effectively one-dimensional. In addition to chemical reactions networks, first-order phase transitions have been considered in

nonequilibrium Ising-like models [91, 94–96, 99]. For such systems, the mean rate of entropy production was also found to be discontinuous at the bistable point, however, the behavior of entropy current fluctuations was not discussed in these works. We expect that by using a similar coarse-graining procedure, the system could also be reduced to a two-state model. Nevertheless, it remains to be seen whether an analytical expression for the coarse-grained transition rates can be obtained for nonequilibrium Ising models.

As a future direction, it would be worthwhile to investigate phase transitions in chemical reactions networks with spatial dynamics [97, 100]. In these systems, the chemical reactions occur locally and diffusion spreads the chemical species. These reaction-diffusion models display spontaneous pattern formation and were first introduced by Turing in 1952 [159]. The behavior of the entropy production at phase transitions remains an open question. Surprisingly, the total entropy production in such systems remains analytical at the phase transition. By splitting the entropy production into a reactive and diffusive part, Falasco *et al.* showed that these contributions have a non-analytical behavior at criticality [97].

We have considered biochemical oscillators to study second-order phase transitions in Chapter 4. Contrary to first-order phase transitions, we found that the fluctuation of entropy current diverges as a power-law at the critical point. We also discussed two criteria for the precision of biochemical oscillations: the number of coherent oscillations and the Fano factor associated with the thermodynamic flux. We initially thought that the Fano factor was an appropriate quantifier for biochemical oscillations as it is a natural observable to quantify precision in light of the thermodynamic uncertainty relation. However, we found that the Fano factor can indicate high precision even if there are no biochemical oscillations. Furthermore, we found that the Fano factor and the number of coherent oscillations can both increase. To summarize, the precision of the thermodynamic flux is not correlated with the precision of oscillations. It turns out that most biochemical oscillators underperform the optimum set by the thermodynamic uncertainty relation [133]. Marsland *et al.* found that a biochemical oscillator can approach this optimum if it has a large number of molecules and internal states.

Biochemical oscillations often occur under the influence of a periodic signal. For instance, circadian clocks rely on the day/night cycle to remain synchronized [119]. Biochemical oscillators must be robust in the presence of strong noise and also sensitive to external cues, which may seem incompatible. Fei *et al.* showed that energy dissipation can increase the entrainment efficiency as well as the precision of oscillations. Periodically driven systems can achieve higher precision at a lower cost compared to autonomous systems [160]. Furthermore, these systems can, in principle, remain coherent forever. This is in contrast to autonomous systems where the number of coherent oscillations has an upper bound [123]. Analogously, Koyuk and Seifert have recently shown that dissipation-less precision can be achieved in periodically driven systems [161]. To summarize, biochemical oscillators can grealty benefit from an external periodic signal to

achieve better precision.

Finally, we have recently proposed an additional way to improve the precision of oscillations. In this thesis, we have assumed that our system is in contact with infinitely large reservoirs that supply chemical species at a constant concentration. In reality, cells are finite systems with a finite number of molecules. As a consequence, real biological oscillators do not have infinitely large reservoirs at their disposal. In [162] we have shown that a system with finite reservoirs can outperform one with ideal reservoirs despite having larger fluctuations. The optimal precision is achieved at finite reservoir size and thermodynamic cost, thus, fluctuations can have a positive effect on the precision of oscillations.

Bibliography

[1] S. Carnot, *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* (Bachelier, Paris, 1824).

[2] L. Onsager, "Reciprocal relations in irreversible processes. I.", Phys. Rev. **37**, 405–426 (1931).

[3] L. Onsager, "Reciprocal relations in irreversible processes. II.", Phys. Rev. **38**, 2265–2279 (1931).

[4] I. Prigogine, *Étude thermodynamique des phénomènes irréversibles* (Desoer, Liège, 1947).

[5] K. Sekimoto, *Stochastic energetics*, Lecture notes in physics (Springer-Verlag, Berlin, Heidelberg, 2010).

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

[7] U. Seifert, "From stochastic thermodynamics to thermodynamic inference", Annu. Rev. Condens. Matter Phys. **10**, 171–192 (2019).

[8] J. Howard, *Mechanics of motor proteins and the cytoskeleton*, 1st ed. (Sinauer, New York, 2001).

[9] M. Schliwa, *Molecular motors* (Wiley-VCH, Weinheim, 2003).

[10] R. Philips, J. Kondev, and J. Theriot, *Physical biology of the cell*, 1st ed. (Garland Science, Taylor & Francis Group, LLC, New York, 2009).

[11] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular biology of the cell*, 6th ed. (Garland Science, Taylor & Francis Group, LLC, New York, 2015).

[12] P. D. Boyer, "The ATP synthase—a splendid molecular machine", Annu. Rev. Biochem. **66**, 717–749 (1997).

[13] W. Junge and N. Nelson, "ATP synthase", Annu. Rev. Biochem. **84**, 631–657 (2015).

[14] S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, "Nonequilibrium energetics of a single F1-ATPase molecule", Phys. Rev. Lett. **104**, 198103 (2010).

[15] S. Toyabe, T. Watanabe-Nakayama, T. Okamoto, S. Kudo, and E. Muneyuki, "Thermodynamic efficiency and mechanochemical coupling of F1-ATPase", Proc. Natl. Acad. Sci. U. S. A. **108**, 17951–17956 (2011).

[16] S. Toyabe and E. Muneyuki, "Single molecule thermodynamics of ATP synthesis by F1-ATPase", New J. Phys. **17**, 015008 (2015).

[17] P. Gaspard and E. Gerritsma, "The stochastic chemomechanics of the F1-ATPase molecular motor", J. Theor. Biol. **247**, 672–686 (2007).

[18] E. Gerritsma and P. Gaspard, "Chemomechanical coupling and stochastic thermodynamics of the F1-ATPase molecular motor with an applied external torque", Biophys. Rev. Lett. **05**, 163–208 (2010).

[19] E. Zimmermann and U. Seifert, "Efficiencies of a molecular motor: a generic hybrid model applied to the F1-ATPase", New J. Phys. **14**, 103023 (2012).

[20] A. C. Barato and U. Seifert, "Thermodynamic uncertainty relation for biomolecular processes", Phys. Rev. Lett. **114**, 158101 (2015).

[21] J. M. Horowitz and T. R. Gingrich, "Proof of the finite-time thermodynamic uncertainty relation for steady-state currents", Phys. Rev. E **96**, 020103 (2017).

[22] P. Pietzonka, A. C. Barato, and U. Seifert, "Universal bound on the efficiency of molecular motors", J. Stat. Mech.: Theory Exp. **2016**, 124004 (2016).

[23] W. Hwang and C. Hyeon, "Energetic costs, precision, and transport efficiency of molecular motors", J. Phys. Chem. Lett. **9**, 513–520 (2018).

[24] M. Campisi and R. Fazio, "The power of a critical heat engine", Nat. Commun. **7**, 11895 (2016).

[25] N. Shiraishi, K. Saito, and H. Tasaki, "Universal trade-off relation between power and efficiency for heat engines", Phys. Rev. Lett. **117**, 190601 (2016).

[26] M. Polettini and M. Esposito, "Carnot efficiency at divergent power output", Europhys. Lett. **118**, 40003 (2017).

[27] P. Pietzonka and U. Seifert, "Universal trade-off between power, efficiency, and constancy in steady-state heat engines", Phys. Rev. Lett. **120**, 190602 (2018).

[28] H. Eugene Stanley, *Introduction to phase transitions and critical phenomena*, 1st ed. (Oxford University Press, Oxford, 1971).

[29] J. M. Yeomans, *Statistical mechanics of phase transitions* (Oxford University Press, Oxford, 1992).

[30] P. Ehrenfest, "Phasenumwandlungen im ueblichen und erweiterten sinn, classifiziert nach den entsprechenden singularitaeten des thermodynamischen potentiales", Proc. Royal Acad. Amsterdam **36**, (Commun. Kamerlingh Onnes Inst. Leiden, Suppl. 75b), 153–157 (1933).

[31] W. H. Keesom and K. Clusiys, "Ueber die specifische wärme des flüssigen helium", Proc. Royal Acad. Amsterdam **35**, (Commun. Kamerlingh Onnes Inst. Leiden, 219e), 307–1320 (1932).

[32] L. Onsager, "Crystal statistics. I. a two-dimensional model with an order-disorder transition", Phys. Rev. **65**, 117–149 (1944).

[33] K. R. Atkins and M. H. Edwards, "Coefficient of expansion of liquid helium ii", Phys. Rev. **97**, 1429–1434 (1955).

[34] G. Jaeger, "The Ehrenfest classification of phase transitions: introduction and evolution", Arch Hist Exact Sc. 53 **53**, 51–81 (1998).

[35] G. Nicolis and M. Malek-Mansour, "Nonequilibrium phase transitions and chemical reactions", Suppl. Prog. Theor. Phys. **64**, 249–268 (1978).

[36] R. Schranner, S. Grossmann, and P. H. Richter, "Dynamic correlation functions in the Brusselator", Z. Physik B **35**, 363–381 (1979).

[37] D. Walgraef, G. Dewel, and P. Borckmans, "Chemical waves in a two-dimensional oscillating system", J. Chem. Phys. **78**, 3043–3051 (1983).

[38] F. Baras, M. M. Mansour, and C. Van den Broeck, "Asymptotic properties of coupled nonlinear Langevin equations in the limit of weak noise. ii: transition to a limit cycle", J. Stat. Phys. **28**, 577–587 (1982).

[39] G. Nicolis, "Dissipative systems", Rep. Prog. Phys. **49**, 873 (1986).

[40] H. Haken, *Synergetics* (Springer-Verlag, Berlin, Heidelberg, 1983).

[41] S. H. Strogatz, ed., *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*, 1st ed. (Cambridge: Westview Press, 2000).

[42] E. Mizuno-Yamasaki, F. Rivera-Molina, and P. Novick, "GTPase networks in membrane traffic", Annu. Rev. Biochem. **81**, 637–659 (2012).

[43] Y. Suda, K. Kurokawa, R. Hirata, and A. Nakano, "Rab GAP cascade regulates dynamics of Ypt6 in the Golgi traffic", Proc. Natl. Acad. Sci. USA **110**, 18976–18981 (2013).

[44] W. M. Bement, M. Leda, A. M. Moe, A. M. Kita, M. E. Larson, A. E. Golding, C. Pfeuti, K.-C. Su, A. L. Miller, A. B. Goryachev, and G. von Dassow, "Activator-inhibitor coupling between Rho signalling and actin assembly makes the cell cortex an excitable medium", Nat. Cell Biol. **17**, 1471–1483 (2015).

[45] A. Ehrmann, B. Nguyen, and U. Seifert, "Interlinked GTPase cascades provide a motif for both robust switches and oscillators", J. R. Soc. Interface **16**, 20190198 (2019).

[46] E. Fischer-Friedrich, G. Meacci, J. Lutkenhaus, H. Chaté, and K. Kruse, "Intra- and intercellular fluctuations in Min-protein dynamics decrease with cell length", Proc. Natl. Acad. Sci. USA **107**, 6134–6139 (2010).

[47] J. Halatek and E. Frey, "Highly canalized MinD transfer and MinE sequestration explain the origin of robust MinCDE-protein dynamics", Cell. Rep. **1**, 741–752 (2012).

[48] L. Xiong and G. Lan, "An optimal free energy dissipation strategy of the MinCDE oscillator in regulating symmetric bacterial cell division", PLoS Comput. Biol. **11**, 1–21 (2015).

[49] F. Wu, J. Halatek, M. Reiter, E. Kingma, E. Frey, and C. Dekker, "Multistability and dynamic transitions of intracellular Min protein patterns", Mol. Syst. Biol. **12**, 873 (2016).

[50] J. Denk, S. Kretschmer, J. Halatek, C. Hartl, P. Schwille, and E. Frey, "MinE conformational switching confers robustness on self-organized Min protein patterns", Proc. Natl. Acad. Sci. USA **115**, 4553–4558 (2018).

[51] F. Schlögl, "Chemical reaction models for non-equilibrium phase transitions", Z. Phys. **253**, 147–161 (1972).

[52] H. K. Janssen, "Stochastisches Reaktionsmodell für einen Nichtgleichgewichts-Phasenübergang", Zeitschrift für Physik **270**, 67–73 (1974).

[53] K. J. McNeil and D. F. Walls, "Nonequilibrium phase transitions in chemical reactions", J. Stat. Phys. **10**, 439–448 (1974).

[54] I. Matheson, D. F. Walls, and C. W. Gardiner, "Stochastic models of firstorder nonequilibrium phase transitions in chemical reactions", J. Stat. Phys. **12**, 21–34 (1975).

[55] J. Nicolis and I. Prigogine, *Self-organization in nonequilibrium systems* (Wiley & Sons, New York, 1977).

[56] I. Prigogine, "Time, structure, and fluctuations", Science **201**, 777–785 (1978).

[57] U. Seifert, "Stochastic thermodynamics: from principles to the cost of precision", Physica A **504**, 176–191 (2018).

[58] T. Schmiedl and U. Seifert, "Stochastic thermodynamics of chemical reaction networks", J. Chem. Phys. **126**, 044101 (2007).

[59] U. Seifert, "Stochastic thermodynamics of single enzymes and molecular motors", English, Eur. Phys. J. E **34**, 26 (2011).

[60] R. Rao and M. Esposito, "Conservation laws shape dissipation", New J. Phys. **20**, 023007 (2018).

[61] R. B. Bapat and T. E. S. Raghavan, *Nonnegative matrices and applications*, Encyclopedia of Mathematics and its Applications (Cambridge University Press, Cambridge, 1997).

[62] K. Sekimoto, "Langevin equation and thermodynamics", Prog. Theor. Phys. Suppl. **130**, 17–27 (1998).

[63] U. Seifert, "Entropy production along a stochastic trajectory and an integral fluctuation theorem", Phys. Rev. Lett. **95**, 040602 (2005).

[64] J. Schnakenberg, "Network theory of microscopic and macroscopic behavior of master equation systems", Rev. Mod. Phys. **48**, 571–585 (1976).

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

[66] H. Eyring, "The activated complex in chemical reactions", J. Chem. Phys. **3**, 107–115 (1935).

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

[68] F. Bouchet and J. Reygner, "Generalisation of the Eyring-Kramers transition rate formula to irreversible diffusion processes", Ann. Henri Poincaré **17**, 3499–3532 (2016).

[69] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", J. Comput. Phys. **22**, 403–434 (1976).

[70] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions", J. Phys. Chem. **81**, 2340–2361 (1977).

[71] D. T. Gillespie, A. Hellander, and L. R. Petzold, "Perspective: stochastic algorithms for chemical kinetics", J. Chem. Phys. **138**, 170901 (2013).

[72] D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems", J. Chem. Phys. **115**, 1716–1733 (2001).

[73] D. T. Gillespie, "Deterministic limit of stochastic chemical kinetics", J. Phys. Chem. B **113**, 1640–1644 (2009).

[74] D. T. Gillespie, "The multivariate Langevin and Fokker–Planck equations", Am. J. Phys. **64**, 1246–1257 (1996).

[75] T. G. Kurtz, "Limit theorems and diffusion approximations for density dependent Markov chains", *Stochastic systems: modeling, identification and optimization, i*, edited by R. J.-. B. Wets (Springer-Verlag, Berlin, Heidelberg, 1976), 67–78.

[76] T. G. Kurtz, "Strong approximation theorems for density dependent Markov chains", Stoch. Proc. Appl. **6**, 223–240 (1978).

[77] C. W. Gardiner, *Handbook of stochastic methods: for physics, chemistry and the natural sciences* (Springer-Verlag, Berlin, Heidelberg, 2004).

[78] N. G. v. Kampen, "A power series expansion of the master equation", Can. J. Phys. **39**, 551–567 (1961).

[79] N. G. van Kampen, *Stochastic processes in physics and chemistry*, 3rd ed., North-Holland Personal Library (Elsevier, Amsterdam, 2007).

[80] E. W. J. Wallace, D. T. Gillespie, K. R. Sanft, and L. R. Petzold, "Linear noise approximation is valid over limited times for any chemical system that is sufficiently large", IET Syst. Biol. **6**, 102–115 (2012).

[81] D. T. Gillespie and E. Seitaridou, eds., *Simple Brownian diffusion: an introduction to the standard theoretical models* (Oxford University Press, Oxford, 2012).

[82] T. Wilhelm, "Chemical systems consisting only of elementary steps - a paradigma for nonlinear behavior", J. Math. Chem. **27**, 71–88 (2000).

[83] M. Vellela and H. Qian, "A quasistationary analysis of a stochastic chemical reaction: Keizer's paradox", Bull. of Math. Biol. **69**, 1727–1746 (2007).

[84] M. Vellela and H. Qian, "Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlögl model revisite", J. R. Soc. Interface **6**, 925 (2009).

[85] H. Ge and H. Qian, "Thermodynamic limit of a nonequilibrium steady state: Maxwell-type construction for a bistable biochemical system", Phys. Rev. Lett. **103**, - (2009).

[86] H. Ge and H. Qian, "Non-equilibrium phase transition in mesoscopic biochemical systems: from stochastic to nonlinear dynamics and beyond", J. R. Soc. Interface **8**, 107–116 (2011).

[87] P. Hänggi, H. Grabert, P. Talkner, and H. Thomas, "Bistable systems: master equation versus Fokker-Planck modeling", Phys. Rev. A **29**, 371–378 (1984).

[88] R. Hinch and S. J. Chapman, "Exponentially slow transitions on a Markov chain: the frequency of calcium sparks", Eur. J. Appl. Math. **16**, 427–446 (2005).

[89] W. J. Heuett and H. Qian, "Grand canonical Markov model: a stochastic theory for open nonequilibrium biochemical networks", J. Chem. Phys. **124**, 044110 (2006).

[90] T. J. Xiao, Z. Hou, and H. Xin, "Entropy production and fluctuation theorem along a stochastic limit cycle", J. Chem. Phys. **129**, 114506 (2008).

[91] L. Crochik and T. Tomé, "Entropy production in the majority-vote model", Phys. Rev. E **72**, 057103 (2005).

[92] B. Andrae, J. Cremer, T. Reichenbach, and E. Frey, "Entropy production of cyclic population dynamics", Phys. Rev. Lett. **104**, 218102 (2010).

[93] T. Rao, T. Xiao, and Z. Hou, "Entropy production in a mesoscopic chemical reaction system with oscillatory and excitable dynamics", J. Chem. Phys. **134**, 214112 (2011).

[94] A. C. Barato and H. Hinrichsen, "Entropy production of a bound nonequilibrium interface", J. Phys. A: Math. Theor. **45**, 115005 (2012).

[95] T. Tomé and M. J. de Oliveira, "Entropy production in nonequilibrium systems at stationary states", Phys. Rev. Lett. **108**, 020601 (2012).

[96] Y. Zhang and A. C. Barato, "Critical behavior of entropy production and learning rate: Ising model with an oscillating field", J. Stat. Mech. **2016**, 113207 (2016).

[97] G. Falasco, R. Rao, and M. Esposito, "Information thermodynamics of Turing patterns", Phys. Rev. Lett. **121**, 108301 (2018).

[98] B. Nguyen, U. Seifert, and A. C. Barato, "Phase transition in thermodynamically consistent biochemical oscillators", J. Chem. Phys. **149**, 045101 (2018).

[99] C. E. F. Noa, P. E. Harunari, M. J. de Oliveira, and C. E. Fiore, "Entropy production as a tool for characterizing nonequilibrium phase transitions", Phys. Rev. E **100**, 012104 (2019).

[100] S. Rana and A. C. Barato, "Precision and dissipation of a stochastic Turing pattern", arXiv 2004.01230 (2020).

[101] Z. Koza, "General technique of calculating the drift velocity and diffusion coefficient in arbitrary periodic systems", J. Phys. A: Math. Gen. **32**, 7637 (1999).

[102] H. Touchette, "The large deviation approach to statistical mechanics", Phys. Rep. **478**, 1–69 (2009).

[103] H. Touchette, "Introduction to dynamical large deviations of Markov processes", Physica A **504**, 5–19 (2018).

[104] M. Uhl and U. Seifert, "Force-dependent diffusion coefficient of molecular brownian ratchets", Phys. Rev. E **98**, 022402 (2018).

[105] J. Keizer, "Nonequilibrium thermodynamics and the stability of states far from equilibrium", Acc. Chem. Res. **12**, 243–249 (1979).

[106] P. Pietzonka, A. C. Barato, and U. Seifert, "Universal bounds on current fluctuations", Phys. Rev. E **93**, 052145 (2016).

[107] H. Grabert and M. S. Green, "Fluctuations and nonlinear irreversible processes", Phys. Rev. A **19**, 1747–1756 (1979).

[108] H. Grabert, P. Hänggi, and I. Oppenheim, "Fluctuations in reversible chemical reactions", Physica A **117**, 300–316 (1983).

[109] H. Qian, "Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems—an analytical theory", Nonlinearity **24**, R19–R49 (2011).

[110] D. Zhou and H. Qian, "Fixation, transient landscape, and diffusion dilemma in stochastic evolutionary game dynamics", Phys. Rev. E **84**, 031907 (2011).

[111] R. Chetrite and H. Touchette, "Nonequilibrium Markov processes conditioned on large deviations", Ann. Henri Poincaré **16**, 2005–2057 (2015).

[112] P. Hänggi, "Nonlinear fluctuations: the problem of deterministic limit and reconstruction of stochastic dynamics", Phys. Rev. A **25**, 1130–1136 (1982).

[113] H. Qian, P. Ao, Y. Tu, and J. Wang, "A framework towards understanding mesoscopic phenomena: emergent unpredictability, symmetry breaking and dynamics across scales", Chem. Phys. Lett. **665**, 153–161 (2016).

[114] J. M. Horowitz, "Diffusion approximations to the chemical master equation only have a consistent stochastic thermodynamics at chemical equilibrium", J. Chem. Phys. **143**, 044111 (2015).

[115] J. L. Lebowitz and H. Spohn, "A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics", J. Stat. Phys. **95**, 333–365 (1999).

[116] M. Van Dyke, *Perturbation methods in fluid dynamics* (Academic Press, New York, 1964).

[117] M. J. Ward, "Exponential asymptotics and convection-diffusion-reaction models", Proceedings of Symposia in Applied Mathematics, AMS Short Courses **56**, 151–184 (1998).

[118] B. Novak and J. J. Tyson, "Design principles of biochemical oscillators", Nat Rev Mol Cell Biol **9**, 981–991 (2008).

[119] A. Goldbeter and M. J. Berridge, *Biochemical oscillations and cellular rhythms: the molecular bases of periodic and chaotic behaviour* (Cambridge University Press, 1996).

[120] M. Nakajima, K. Imai, H. Ito, T. Nishiwaki, Y. Murayama, H. Iwasaki, T. Oyama, and T. Kondo, "Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro", Science **308**, 414–415 (2005).

[121] J. E. Ferrell, T. Y.-C. Tsai, and Q. Yang, "Modeling the cell cycle: why do certain circuits oscillate?", Cell **144**, 874–885 (2011).

[122] Y. Cao, H. Wang, Q. Ouyang, and Y. Tu, "The free-energy cost of accurate biochemical oscillations", Nat Phys **11**, 772–778 (2015).

[123] A. C. Barato and U. Seifert, "Coherence of biochemical oscillations is bounded by driving force and network topology", Phys. Rev. E **95**, 062409 (2017).

[124] H. Wierenga, P. R. ten Wolde, and N. B. Becker, "Quantifying fluctuations in reversible enzymatic cycles and clocks", Phys. Rev. E **97**, 042404 (2018).

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

[126] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, "The energy-speed-accuracy trade-off in sensory adaptation", Nat. Phys. **8**, 422–428 (2012).

[127] P. Mehta and D. J. Schwab, "Energetic costs of cellular computation", Proc. Natl. Acad. Sci. USA **109**, 17978–17982 (2012).

[128] C. C. Govern and P. R. ten Wolde, "Energy dissipation and noise correlations in biochemical sensing", Phys. Rev. Lett. **113**, 258102 (2014).

[129] D. Hartich, A. C. Barato, and U. Seifert, "Nonequilibrium sensing and its analogy to kinetic proofreading", New J. Phys. **17**, 055026 (2015).

[130] T. E. Ouldridge, C. C. Govern, and P. R. ten Wolde, "Thermodynamics of computational copying in biochemical systems", Phys. Rev. X **7**, 021004 (2017).

[131] C. Fei, Y. Cao, Q. Ouyang, and Y. Tu, "Design principles for enhancing phase sensitivity and suppressing phase fluctuations simultaneously in biochemical oscillatory systems", Nat. Commun. **9**, 1434 (2018).

[132] T. Herpich, J. Thingna, and M. Esposito, "Collective power: minimal model for thermodynamics of nonequilibrium phase transitions", Phys. Rev. X **8**, 031056 (2018).

[133] R. Marsland, W. Cui, and J. M. Horowitz, "The thermodynamic uncertainty relation in biochemical oscillations", J. R. Soc. Interface **16**, 20190098 (2019).

[134] C. del Junco and S. Vaikuntanathan, "High chemical affinity increases the robustness of biochemical oscillations", Phys. Rev. E **101**, 012410 (2020).

[135] C. del Junco and S. Vaikuntanathan, "Robust oscillations in multi-cyclic Markov state models of biochemical clocks", J. Chem. Phys. **152**, 055101 (2020).

[136] D. Zhang, Y. Cao, Q. Ouyang, and Y. Tu, "The energy cost and optimal design for synchronization of coupled molecular oscillators", Nat. Phys. **16**, 95–100 (2020).

[137] M. Schnitzer and S. Block, "Statistical kinetics of processive enzymes", Cold Spring Harbor Symp. Quant. Biol. **60**, 793 (1995).

[138] Y. R. Chemla, J. R. Moffitt, and C. Bustamante, "Exact solutions for kinetic models of macromolecular dynamics", J. Phys. Chem. B **112**, 6025–6044 (2008).

[139] J. R. Moffitt and C. Bustamante, "Extracting signal from noise: kinetic mechanisms from a Michaelis–Menten-like expression for enzymatic fluctuations", FEBS Journal **281**, 498–517 (2014).

[140] A. C. Barato and U. Seifert, "Universal bound on the Fano factor in enzyme kinetics", J. Phys. Chem. B **119**, 6555–6561 (2015).

[141] A. C. Barato and U. Seifert, "Skewness and kurtosis in statistical kinetics", Phys. Rev. Lett. **115**, 188103 (2015).

[142] A. C. Barato and U. Seifert, "Cost and precision of brownian clocks", Phys. Rev. X **6**, 041053 (2016).

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

[144] H. Qian and M. Qian, "Pumped biochemical reactions, nonequilibrium circulation, and stochastic resonance", Phys. Rev. Lett. **84**, 2271–2274 (2000).

[145] P. Gaspard, "The correlation time of mesoscopic chemical clocks", J. Chem. Phys. **117**, 8905–8916 (2002).

[146] Z. Hou, T. J. Xiao, and H. Xin, "Internal noise coherent resonance for mesoscopic chemical oscillations: a fundamental study", ChemPhysChem **7**, 1520–1524 (2006).

[147] T. Xiao, J. Ma, Z. Hou, and H. Xin, "Effects of internal noise in mesoscopic chemical systems near Hopf bifurcation", New J. Phys. **9**, 403 (2007).

[148] L. G. Morelli and F. Jülicher, "Precision of genetic oscillators and clocks", Phys. Rev. Lett. **98**, 228101 (2007).

[149] D. J. Jörg, L. G. Morelli, and F. Jülicher, "Chemical event chain model of coupled genetic oscillators", Phys. Rev. E **97**, 032409 (2018).

[150] J. S. van Zon, D. K. Lubensky, P. R. H. Altena, and P. R. ten Wolde, "An allosteric model of circadian KaiC phosphorylation", Proc. Natl. Acad. Sci. USA **104**, 7420–7425 (2007).

[151] G. Dong and S. S. Golden, "How a cyanobacterium tells time", Curr. Opin. Microbiol. **11**, 541–546 (2008).

[152] R. Lefever, G. Nicolis, and P. Borckmans, "The Brusselator: it does oscillate all the same", J. Chem. Soc., Faraday Trans. 1 **84**, 1013–1023 (1988).

[153] H. Qian, S. Saffarian, and E. L. Elson, "Concentration fluctuations in a mesoscopic oscillating chemical reaction system", Proc. Natl. Acad. Sci. USA **99**, 10376–10381 (2002).

[154] D. Andrieux and P. Gaspard, "Fluctuation theorem and mesoscopic chemical clocks", J. Chem. Phys. **128**, 154506 (2008).

[155] J. Paijmans, D. K. Lubensky, and P. R. ten Wolde, "Period robustness and entrainability of the Kai system to changing nucleotide concentrations", Biophys. J. **113**, 157–173 (2017).

[156] A. J. McKane, J. D. Nagy, T. J. Newman, and M. O. Stefanini, "Amplified biochemical oscillations in cellular systems", J. Stat. Phys. **128**, 165–191 (2007).

[157] M. J. de Oliveira, "Irreversible models with Boltzmann–Gibbs probability distribution and entropy production", J. Stat. Mech.: Theory Exp. **2011**, P12012 (2011).

[158] G. Nicolis and J. Turner, "Stochastic analysis of a nonequilibrium phase transition: some exact results", Physica A **89**, 326–338 (1977).

[159] A. M. Turing, "The chemical basis of morphogenesis", Philos. Trans. R. Soc., B **237**, 37–72 (1952).

[160] L. Oberreiter, U. Seifert, and A. C. Barato, "Subharmonic oscillations in stochastic systems under periodic driving", Phys. Rev. E **100**, 012135 (2019).

[161] T. Koyuk and U. Seifert, "Operationally accessible bounds on fluctuations and entropy production in periodically driven systems", Phys. Rev. Lett. **122**, 230601 (2019).

[162] J. H. Fritz, B. Nguyen, and U. Seifert, "Stochastic thermodynamics of chemical reactions coupled to finite reservoirs: a case study for the Brusselator", J. Chem. Phys. **152**, 235101 (2020).

Danksagung

Als erstes gilt mein Dank Herrn. **Prof. Dr. Udo Seifert** für die Betreuung während meiner Promotion, nicht nur die Freiheit in der Richtung meiner Forschung, sondern auch für wertvolle Ratschläge und schließlich für die Gelegenheit, ihn bei seinen Vorlesungen mehrmals zu vertreten.

Herr **Prof. Dr. Andre Cardoso Barato** danke ich für die erflogreiche Zusammenarbeit, aus der ein Kapitel der hier präsentierten Ergebnisse enstand. Herr **Prof. Dr. Siegfried Dietrich** danke ich für die Übernahme des Mitberichts. Besonderer Dank gilt Herr **Dr. David Hartich**, der ein exzellenter Mentor während meiner ersten Monate war und mich motiviert hat, an diesem Institut zu promovieren. Frau **Anja Steinhauser** danke ich herzlich für die freundliche und professionelle Unterstützung bei allen organisatorischen Belangen.

Allen aktuellen und ehemaligen Institutsmitgliedern danke ich für die gemütliche Arbeitsatmosphäre, die vielen spannenden und unterhaltsamen Diskussionen. Für zahlreiche Diskussionen über die Theorie der großen Abweichungen und den Diffusionskoeffizienten möchte ich **Lukas Fischer** und **Matthias Uhl** herzlich danken. Ein besonderer Dank geht an unser Admin-Team am Institut, dem es immer gelang, für jedes Problem eine schnelle und effiziente Lösung zu finden.

Außerdem möchte ich allen höchst engagierten Studierenden danken, die ich während meiner Promotion bei einer erfolgreichen Abschlussarbeiten betreuen durfte. Darunter sind die Masterarbeit von **Andreas Ehrmann** und die Bachelorarbeiten von **Julius Degünther** und **Jonas Fritz** zu nennen, die es mir ermöglichten, spannende Themen zu erforschen. Vielen Dank an euch drei für die Erfahrung.

Vielen Dank an Matthias Uhl und Julius Degünther für das inhaltliche und sprachliche Korrekturlesen dieser Arbeit.

Abschließend möchte ich mich bei meiner gesamten Familie für die Unterstützung und den Rückhalt während des gesamten Studiums bedanken.

Ehrenwörtliche Erklärung

Ich erkläre, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Petit-Lancy (Schweiz), den XX. Mai 2020 Basile Nguyen