Limits for Stochastic Reaction Networks

Limits for Stochastic Reaction Networks

PhD thesis by

DANIELE CAPPELLETTI

Department of Mathematical Sciences
University of Copenhagen
Denmark

PhD School of Science - Faculty of Science - University of Copenhagen

Daniele Cappelletti
Department of Mathematical Sciences
University of Copenhagen
Universitetsparken 5
DK-2100 København Ø
Denmark

d.cappelletti@math.ku.dk
htpp://www.math.ku.dk/∼d.cappelletti

PhD thesis submitted to the PhD School of Science, Faculty of Science, University of Copenhagen, Denmark, in October 2015.

Academic advisor: Carsten Wiuf
University of Copenhagen, Denmark

Assessment Committee: Niels Richard Hansen (chair)
University of Copenhagen, Denmark

Peter Pfaffelhuber
University of Freiburg, Germany

János Tóth
Budapest University of Technology and Economics, Hungary

ISBN 978-87-7078-937-0

Preface

This dissertation is submitted in partial fulfillment of the requirements for the Ph.D. degree at the Faculty of Science, University of Copenhagen, Denmark. The work was carried out at the Department of Mathematical Sciences, University of Copenhagen, from October 2012 to September 2015. It was financed by the Department of Mathematical Sciences, the Carlsberg Foundation, the Lundbeck Foundation and the Danish Cancer Society.

Paper I and Paper II are different form those in the original version of this thesis. Specifically, in the original version of Paper I a minor error was present, which required an assumption (cfr. equation (4.9)) to be slightly changed. Proposition 4.1 has been added to connect the new assumption with the original one. We thank the student Peter Jenni from University of Freiburg for his help with spotting the error. Paper II has been updated after receiving reviews from the journal it was submitted to. Essentially, nothing has changed except for a wider definition of *complex balanced equilibria* that are now allowed to be non-negative vectors. Consequently, Theorem 3.1 has been stated in a more general formulation, and a proof has been added for it in Appendix B. Finally, calculations regarding Examples 2 and 4 have been added in Appendix C.

This work is mostly the result of the fruitful collaboration with my supervisor, Carsten Wiuf. Moreover, part of this work would not have been in the present form without further precious comments by Elisenda Feliu. Work with other members of the Statistics and Probability Theory Research Group at the Department of Mathematical Science have also been important, as well as getting involved and well accepted in the Chemical Reaction Network international community. I would like to thank Carsten Wiuf and Elisenda Feliu for introducing me to the community and for their guidance. Finally, a pleasant and meaningful collaboration took place in Madison during my three month stay abroad. I am grateful to David Anderson, Gheorghe Craciun and Thomas Kurtz for their warm hospitality and for the significant discussions we had. Finally, I would like to thank the Department of Mathematical Science for supporting my travels, which have been fundamental for me to get involved in the international framework of chemical reaction network theory.

Summary

Reaction systems have been introduced in the 70s to model biochemical systems. Nowadays their range of applications has increased and they are fruitfully used in different fields. The concept is simple: some chemical species react, the set of chemical reactions form a graph and a rate function is associated with each reaction. Such functions describe the speed of the different reactions, or their propensities. Two modelling regimes are then available: the evolution of the different species concentrations can be deterministically modelled through a system of ODE, while the counts of the different species at a certain time are stochastically modelled by means of a continuous-time Markov chain. Our work concerns primarily stochastic reaction systems, and their asymptotic properties.

In Paper I, we consider a reaction system with intermediate species, i.e. species that are produced and fast degraded along a path of reactions. Let the rates of degradation of the intermediate species be functions of a parameter N that tends to infinity. We consider a reduced system where the intermediate species have been eliminated, and find conditions on the degradation rate of the intermediates such that the behaviour of the reduced network tends to that of the original one. In particular, we prove a uniform punctual convergence in distribution and weak convergence of the integrals of continuous functions along the paths of the two models. Under some extra conditions, we also prove weak convergence of the two processes. The result is stated in the setting of multiscale reaction systems: the amounts of all the species and the rates of all the reactions of the original model can scale as powers of N. A similar result also holds for the deterministic case, as shown in Appendix IA.

In Paper II, we focus on the stationary distributions of the stochastic reaction systems. Specifically, we build a theory for stochastic reaction systems that is parallel to the deficiency zero theory for deterministic systems, which dates back to the 70s. A deficiency theory for stochastic reaction systems was missing, and few results connecting deficiency and stochastic reaction systems were known. The theory we build connects special form of product-form stationary distributions with structural properties of the reaction graph of the system.

In Paper III, a special class of reaction systems is considered, namely systems

exhibiting *absolute concentration robust* species. Such species, in the deterministic modelling regime, assume always the same value at any positive steady state. In the stochastic setting, we prove that, if the initial condition is a point in the basin of attraction of a positive steady state of the corresponding deterministic model and tends to infinity, then up to a fixed time T the counts of the species exhibiting absolute concentration robustness are, on average, near to their equilibrium value. The result is not obvious because when the counts of some species tend to infinity, so do some rate functions, and the study of the system may become hard. Moreover, the result states a substantial concordance between the paths of the stochastic and the deterministic models.

Short abstract

Reaction systems are mathematical models used in biochemistry and in a number of other fields. They model the evolution of a biochemical mechanism either deterministically, by means of a system of ODEs, or stochastically, by means of a continuous-time Markov chain. Our main concern is the stochastic modelling regime: we study asymptotic results for the associated Markov chain when some parameters of the model tend to infinity, in order to approximate the model by a simpler one or to recover some features from its deterministic counterpart. We also study the stationary distributions of the model. Specifically, in the first paper collected in this manuscript the original stochastic model is approximated by a lower dimensional one, where particular chemical species have been eliminated. We study different kind of convergence of the reduced model to the original one, and we also prove a similar result for the deterministic model. In the second paper connections between the form of the stationary distribution of a stochastic system and structural conditions of the underlying chemical reactions are unveiled. Finally, in the third paper we prove that if a chemical species has always the same value for any positive steady state of a deterministic reaction system, then the counts of that species in the stochastic model are, on average and up to a finite time T, near to that value. This results holds when the counts of the other species in the initial condition tend to infinity: in such a situation the production and degradation rates tend to infinity as well, and the evolution of the different species counts is not clear a priori.

Dansk resumé

Reaktionssystemer er matematiske modeller, der blandt andet anvendes indenfor biokemi og en række andre videnskabelige områder. Disse systemer modellerer udviklingen af en biokemisk mekanisme enten deterministisk ved hjælp af et system af almindelige differential ligninger eller stokastisk ved hjælp af en kontinuert tids Markov kæde. Vores primære interesse er stokastisk modellering. Specielt udleder vi asymptotiske resultater vedrørende den Markov kæde, der beskriver systemet, når nogle parametre eller variable i modellen er store ('går mod uendelig'). Dette er dels med henblik på at beskrive modellen med en simplere model (der fremkommer under de asymptotiske betingelser) og dels med henblik på at sammenholde den deterministiske model til den stokastiske. Desuden har vi studeret stationære fordelinger for de stokastiske modeller baseret på kontinuert tids Markov kæder.

I den første artikel i denne afhandling tilnærmes den oprindelige stokastiske process med en process af lavere dimension, der opnås ved at 'eliminere' visse kemiske stoffer (variable) fra modellen. Vi studerer forskellige typer af konvergens af den oprindelige model til den reducerede model. Desuden viser vi et tilsvarende resultat for deterministiske modeller baseret på differentialligninger. I den anden artikel knyttes forbindelser mellem den stationære fordeling for en stokastisk model og ligevægtspunkter af en tilsvarende deterministiske model. Vi viser at strukturelle betingelser der er formuleret for deterministiske modeller og som har stærke implikationer for karakteriseringen af ligevægtspunkterne, har en analog stokastisk formulering. Endelig i den tredje artikel beviser vi en relation mellem deterministiske og stokastiske systemer for en bestemt klasse af reaktionssystemer: Hvis et kemisk stof (variable) i et deterministisk system har den samme værdi i ligevægt for enhver startbetingelse, så er middelværdien af dette stof til enhver endelig tid i det stokastiske system tæt på den deterministiske ligevægtsværdi. Dette resultat gælder under visse betingelser på reaktionshastigheder og på antallet af molekyler af de andre kemiske stoffer i systemet.

Contents

Papers

1 Introduction

1.1 Informal description of the models

Mathematics has been intensively used for modelling biochemical systems. In particular, the behaviour of systems of interacting species is usually studied by means of *reaction networks*. These are particular mathematical structures that can give rise to distinct dynamical models, known as deterministic and stochastic reaction systems (Érdi & Tóth, 1989; Anderson & Kurtz, 2011, 2015). Specifically, a reaction network can be thought as the collection of the transformations that the species undergo in the biochemical system of interest: an example of transformation could be $A + B \to C$, where A, B and C are chemical species, and the bound between two molecules of A and B may form a molecule of C. The left-hand side and the right-hand side of the transformations are linear combinations of the species, and are referred to as *complexes*. Furthermore, the transformations are named *reactions*, and constitute a directed graph (the *reaction graph*) with the set of complexes as nodes.

The main focus resides in modelling the dynamics of the biochemical system of interest. In order to do this, first let $z(t)$ be a real vector whose entries are non-negative and express the concentration of the different chemical species at a specific time t. A rate can be then associated with each reaction, as an example we can associate the rate $\lambda(z) = \kappa z_A z_B$ with the reaction $A + B \to C$, for a positive constant κ. By considering the consumption and production of species given by each reaction at the given rates, we model the evolution of $z(t)$ by means of an ordinary differential equation. This model is known as *deterministic reaction system*, more details are given in Chapter 2.
Imagine, on the other hand, that few molecules are present in the system and that we are therefore concerned about their counts, rather than their concentrations. Let $X(t)$ be a vector whose entries are natural numbers expressing such counts at a specific time t. In this case we can imagine that for most of the time the molecules float without interacting, and that the time to wait for the next reaction to take place is a random variable, as well as which is the next occurring reaction. When a

reaction takes place, the vector $X(t)$ is changed accordingly. For example, if at time t^* the reaction $A+B \to C$ occurs, then $X(t^*)$ is given by the previous state $X(t^*-)$ where the entries of the species A and B have been decreased by one, and the entry of C is increased by one. The evolution of $X(t)$ is modelled by a continuous-time Markov chain, where the reaction rates are the transition rates. The model is named *stochastic reaction system*, and a more detailed and formal introduction is given in Chapter 2.

In both the deterministic and the stochastic models, a choice of rate functions for a reaction network is called *kinetics*. A popular choice is the *mass-action kinetics*, which corresponds to the hypothesis that the chemical components are well stirred. Such kinetics assigns a polynomial rate function to each reaction, which is proportional to the ways the present molecules can combine to give rise to the reaction. The proportionality constants, termed *rate constants*, uniquely determine the kinetics both in the deterministic and the stochastic models. A formal introduction to mass-action kinetics is given in later chapters. As an example, the mentioned rate function $\lambda(z) = \kappa z_A z_B$ for the reaction $A + B \to C$ corresponds to mass-action kinetics for the deterministic model, and κ is a rate constant. The common choice of polynomial kinetics (not only mass-action) for reaction systems attracted the interest of many algebraists, who gave important contributions to chemical reaction network theory, especially in the deterministic setting (Craciun *et al.*, 2009; Karp *et al.*, 2012; Müller *et al.*, 2013; Feliu, 2015).

1.2 Applications

Despite the formulation of the models we gave in terms of interacting chemical species, chemical reaction network theory (which we will abbreviate by CRNT) can be applied outside the field of biochemistry. Indeed, the symbols used for chemical species may as well denote interacting organisms, or individuals with different characteristics, or any kind of interacting objects. As a matter of fact, the Lotka Volterra equations (Murray, 2002), dating back to 1910, can be seen as an early example of deterministic reaction system where the interaction between a predator A and a prey B is modelled through the reactions $0 \to A$ (birth of a prey), $A + B \to B$ (consumption of a prey), $A + B \to B$ (birth of a predator in abundance of preys) and $B \to 0$ (death of a predator). The considered kinetics was mass-action kinetics.

An early example of stochastic reaction system can be found in the SIR model of diseased spread (Kermack & McKendrick, 1927), even though CRNT was not mentioned. In this model we distinguish between susceptible, infected and recovered individuals denoted by S, I and R respectively. The dynamics are governed by the two reactions $S + I \to 2I$ (infection) and $I \to R$ (recovery). Here, as well, the usual

choice for the transition rates corresponds to mass-action kinetics.

CRNT has been fruitfully applied in a wide range of biological fields, such ecology (May, 2001), epidemiology (Anderson & May, 1992), biochemistry (Cornish-Bowden, 2004), cellular biology (Ingram *et al.*, 2008) and systems biology (Karlebach & Shamir, 2008). As an example outside biology, chemical reaction networks can model queueing systems through reactions like $0 \to C_i$ (arrival of a customer at the queue i), $C_i \to C_j$ (change of queue) and $C_i \to 0$ (departure of a customer from the queue i). In this context, it is not surprising that connections can be found between CRNT and Petri nets (Angeli *et al.*, 2007; Mairesse & Nguyen, 2009). CRNT has been further applied in socio-economy (Peschel & Breitenecker, 1984; Peschel & Mende, 1986) and in quantitative sociology (Weidlich & Haag, 2012). Moreover, an interest in chemical reaction networks as programming language has been recently developed: the power of the language of chemical reaction networks (Rothemund & Winfree, 2000) and the possibility to physically implement them (Soloveichik *et al.*, 2010) arose interest in the computer science field (Doty *et al.*, 2012; Cummings *et al.*, 2014).

In all the fields where CRNT has been applied, reaction systems serve as models for the time evolution of interacting objects. If the models are accurate enough, then they can be used to make predictions on the behaviour of real world phenomena and to design ways to interact with them: well-understood phenomena can be used for constructing new technology (as bio-computers that need no electricity to run), and controlled interventions can be designed (as curing a disease by prescribing drugs, or as intervention in an ecological system in order to save a species from extinction). Designing a model and testing it are certainly the first steps to understand a phenomenon.

The dynamics of reaction systems can be simulated on a computer, which leads to a consistent reduction of the time and the resources needed for experiments. In particular, the outcome of simulations may suggest further investigations or a more meaningful calibration of the experiment parameters: if the simulations are compared to the observed phenomena, then the parameters could be better tuned, or the proposed description of the interactions between the objects of interest could be corrected.

Besides computational simulations, which with no doubt offer an important help in trying to understand the mechanisms of interacting species, direct mathematical analysis of the reaction network can be performed. The latter provides perhaps a more robust insight: given a reaction system, by means of simulations we may understand which kind of evolutions are feasible, while by performing a mathematical analysis we can understand why. Specifically, a typical question in CRNT would be what dynamical properties can be exhibited by a reaction system regardless the choice of kinetics, or at least by only assuming that the rate functions assume a

certain form. Such qualitative study is important because inferring the rate constants from observations is often prohibitive, as is tuning the parameters of an experiments with high accuracy. Concrete examples of qualitative characteristics of interest could be whether a designed bio-chip computes the desired outcome in finite time, whether the suppression of a chemical reaction in a cell could lead to a different form of stationary behaviour, or whether a certain modification in an ecological environment leads to the survival of all the species involved.

Two major concerns of CRNT are the study of the transient and the stationary behaviours of the reaction systems. In the former case, the trajectory in compact intervals of time is studied: for example, this is done in the classical scaling (see Section 2.4.3), where the evolution of a stochastic reaction system is considered up to a fixed time T, under the hypothesis that the counts of the different species are high. On the other hand, the study of stationary behaviours concerns systems at equilibrium or at operating speed, and the long term behaviours of the models. The mathematical objects of interest in this case are steady states and limit cycles for the deterministic reaction systems, and stationary distributions for the stochastic reaction systems. As an example in the setting of stochastic models, imagine that we are interested in designing a bio-chip that calculates the sum of two natural numbers, say $a + b$. As a chip, we can think of a box where a molecules of a species A and b molecules of a species B are introduced, the reactions $A \to C$ and $B \to C$ occur, and the counts of the species C are then considered. It is not hard to see that eventually the counts of the species C will be exactly $a + b$, which is the function that we wanted the chip to calculate. In this case, our main concern is about the final outcome, described by the stationary distribution and the absorbing states of the stochastic system.

In the deterministic setting, the number of steady states is of interest: in a simplified setting, we can consider the steady states as the equilibrium conditions to which the observed phenomenon tends in a short time. In this case, the existence of multiple steady states indicates the ability of the mechanism to actively respond to the change of the environment by switching to one steady state to another. This behaviour is called "biological switch" and it is intensively studied in cellular biology (Markevich *et al.*, 2004; Thomson & Gunawardena, 2009)'. Different mathematical frameworks to recognize the *multistationarity* of a system by structural conditions of the underlying reaction graph have been proposed (Horn & Jackson, 1972; Feinberg, 1987, 1988; Feliu & Wiuf, 2013; Müller *et al.*, 2013; Joshi & Shiu, 2013; Feliu, 2015), though much work has yet to be done in this direction.

In the deterministic setting, different stationary behaviours can be exhibited, which are not captured by the presence of steady states. As an example, trajectories can tend to a limit cycle, and therefore exhibit sustained *oscillations*: this is the case of many important biological mechanisms including the "circadian clock". The

circadian clock is a biochemical mechanism responsible for regulating the sleep cycle in our body, as well as for regulating the response of our body to the changing seasons in the course of a year. The nature of this regulating mechanism at equilibrium is of course cyclic, and cycles of different periods are involved (one day and one year). Possible cures and further understanding of observed malfunctioning of the circadian clock are a topic of great interest in biochemistry, however the study of this system is difficult because of its complexity and because different changes happen on different time scales (the process is a *multiscale* process). This problem sheds light on another important use of CRNT in biochemistry: accurate mathematical analysis of the models can provide means to successfully reduce its complexity in order to study the characteristics of interest more efficiently. It would be computationally prohibitive to simulate very large biochemical systems, and deep mathematical knowledge has to be applied to perform a careful approximation of them. For example, if we want to study the changes of the circadian clock across the different seasons, we might reduce the system by averaging the changes happening at the time scales of minutes and days, while if we want to study the mechanism that interfere with cellular regulations in the scale of minutes, we consider the chemical species whose changes occur on larger time scales constant. Reductions can sometimes unveil what the essential structure of a biological systems is on the different time scales and which reactions are mainly responsible for certain outcomes.

Regimes that are neither transient nor stationary might be of interest as well. This is perhaps better explained with an example. Consider the Michaelis-Menten mechanism

$$E + S \rightleftarrows ES \longrightarrow E + P$$

which is used to model enzymatic transformation of a substrate S into a product P by means of the enzyme E (Cornish-Bowden, 2004). Assume, moreover, that the kinetics is mass-action kinetics. Since enzymatic transformations appear frequently in cellular biology, this simple model has been intensively studied since the beginning of the twentieth century. A common assumption is the presence of a large concentration of substrate and a much less concentration of enzyme. If the abundance of substrate exceeds a certain threshold, the solution is saturated and, after a short period and up to the eventual total consumption of substrate S, at any time point almost all the enzymes E are bound to a substrate molecule in the complex ES. Hence, the rate of the reaction $ES \to E + P$ is practically constant. This behaviour is of interest because it can be observed for a long period, and can be considered as a "quasi-stationary phase" of the system: it is manifested after a short transient phase and precedes the proper stationary phase of the system, which is the total degradation of S. It cannot be considered as an apparent equilibrium of the system because degradation of S and production of P is taking place, however it is a temporary

"dynamical" equilibrium in the sense that the rates of the reactions are substantially unchanged for a long time. It is worth noting that without the presence of an intermediate complex ES, the described quasi-stationary phase would not exist. In fact, the intermediate complex is usually not directly detectable in the experiments and its existence has been suggested by the discrepancies of the observed dynamics of enzymatic mechanisms and those predicted by the simpler model $E + S \to E + P$ (Gunawardena, 2012).

Quasi-stationary phases are of interest whenever the systems reach an apparent dynamical or static equilibrium for a large period of time: in the stochastic setting the latter can be captured by the study of *quasi-stationary distributions*. As an example we can consider a stochastic reaction system in ecology: if the model is accurate enough, then the only possible stationary distribution is concentrated on the extinction of all the species of interest, since every species eventually gets extinct in nature. However, the extinction state is usually hit after a long waiting time, and a quasi-stationary distribution of the model could be of interest if we want to predict the behaviour of the system in the next twenty years.

1.3 My contribution in the international state-of-art

The papers collected in this thesis contain my contribution in the field of CRNT, developed at the Department of Mathematical Sciences, University of Copenhagen, from October 2012 to September 2015. The aim of the articles collected here and their relation to international state-of-the-art will be better explained in Chapter 3, after all the necessary mathematical background and the relevant known results in CRNT have been introduced. However, the main objectives of our investigation can be ultimately put down to complexity reduction of the network and study of the relationship between the deterministic and stochastic models. The focus of the study is both on the transient behaviours (Paper I, Appendix IA and Paper III) and on the stationary distributions (Paper II).

Mechanisms in cellular biology are often very large and therefore analytically and computationally untreatable. Moreover, it is often the case that the abundances of the different reactants and the rates of the different reactions vary across different levels of magnitude, as for the circadian clock. Such reaction systems are called *multiscale systems* and their simulations are even more complicated. Approximation of the dynamics has always been an issue, since the beginning of the study of enzyme kinetics (Cornish-Bowden, 2004), and it has become more compelling for the study of more complex structures. In this setting, Paper I offers a way to approximate the original stochastic reaction system by means of a reduced model, where

special species called *intermediates* are removed. Such species are quickly consumed intermediaries of chemical transformations. For example, in the reaction network $A \to H \to B$, H is an intermediate and, under certain assumptions on the kinetics, we propose the reduction of the network to $A \to B$. In the paper, we study in detail what is and what is not preserved by the reduction, assuming that the consumption rates of the intermediates tend to infinity. Another example of intermediate species is ES in the Michaelis-Menten mechanism. The paper lies in the framework of Kurtz (1972), Kurtz (1978), Ball *et al.* (2006), Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013), which focus on asymptotic results in order to simplify the study of stochastic reaction systems. Kurtz (1972) is perhaps the first paper to connect stochastic and deterministic reaction networks, and this is achieved by showing that the deterministic model is in fact the weak limit of the stochastic model, under certain hypothesis. Kurtz (1978) is a refinement of the results in Kurtz (1972), while Ball *et al.* (2006), Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013) indicate piecewise deterministic stochastic processes as weak limits for multivariate stochastic reaction networks. In the framework of the latter papers, particular species and reactions are eliminated from the original models, however a reduction by means of elimination of intermediates as described in Paper I cannot be performed. On the other hand, the approximating model we introduce can be further simplified through techniques developed in Ball *et al.* (2006), Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013), to which our Paper I can be therefore considered as complementary.

In Appendix IA we show that the very same reduced model as in Paper I serves also in the deterministic setting as approximation of the trajectory of the original model. There is more: in the deterministic setting, algebraic techniques have been succesfully employed to obtain reduced reaction systems. In particular, in Feliu & Wiuf (2013) a reduced reaction system is obtained from the original deterministic reaction system by elimination of intermediates species, and the number and type of steady states of the two systems are related. Though our simplified model is described in a different way and targets approximation of trajectories, in Paper I we are able to prove it coincides with the reduced model from Feliu & Wiuf (2013), showing in that way that the same simplified model serves different purposes.

We will see in later chapters that some dynamical properties of a reaction system (either deterministic or stochastic) may depend on the particular choice of kinetics, while in some special cases they are uniquely determined by the structural properties of the reaction graph. Much effort has been put in investigating the connection between graphical properties of a reaction network and the dynamics of a corresponding reaction system: apart from developing a set of interesting theoretical results, the benefit of such investigation relies on the simplicity of working with graphical conditions, in contrast with the difficulty to manage high-order ODEs or stochastic processes, and the errors that affect rate functions inferred from biological

experiments. We refer to Horn & Jackson (1972), Feinberg (1972) and Horn (1972) as the pioneer works in this sense and as the fundamentals of *deficiency theory*, later explained in this thesis (see also Feinberg, 1979, 1987, 1988). The main goal of this theory is to infer the number and the type of steady states of a deterministic reaction system by simple graphical conditions. Further related work has been recently carried out, for example, in Shinar & Feinberg (2010), Boros (2013) and Craciun (2015). For systems not falling into the setting of these papers, the possibility of multiple steady states generally depends on the kinetics chosen and is not an easy problem to address, but much has been done in this direction (Feliu & Wiuf, 2013; Müller *et al.*, 2013; Joshi & Shiu, 2013; Feliu, 2015). Despite an increasing interest in stochastic reaction systems, few results connect graphical properties of the network with dynamical features of the stochastic model: Jahnke & Huisinga (2007), Anderson *et al.* (2010), Anderson *et al.* (2014a), Anderson *et al.* (2014b), Gupta & Khammash (2014) and Joshi (2015) are nearly the only papers in this direction. The four of them deal with the stationary distribution of the stochastic reaction systems, in different situations. They also study the relationship between deterministic and stochastic models, in the following sense: if a reaction network is considered and a choice of rate constants is made, then we can either decide to consider the associated deterministic mass-action system or the stochastic one. The addressed question is how they are related, specifically how the steady states of the deterministic models are connected to the stationary distribution of the stochastic model. In Anderson *et al.* (2010), existence and uniqueness of an explicit product-form Poisson-like stationary distribution is proven for stochastic reaction systems whose reaction graph has *deficiency zero* and is *weakly reversible*. Such result can be extended with no effort to the broader family of *complex balanced* systems, which will be defined in Chapter 2. In this case, the mean of the stationary distribution is strongly related to the steady state of the deterministic model. In Paper II, we expand the main result of Anderson *et al.* (2010) and fill the gap between the stochastic and deterministic models by building a parallel deficiency theory for stochastic reaction systems. We also define the concept of "complex balancing" in the stochastic setting, and prove that it coincides with the known deterministic counterpart.

In Anderson *et al.* (2014b), a discrepancy between the dynamics of the stochastic and deterministic systems is unveiled. As an example, consider the reaction network

$$A + B \longrightarrow 2B \qquad\qquad B \longrightarrow A$$

and fix two positive rate constants for the reactions. For the associated deterministic mass-action system, any positive steady state exhibits the same concentration for the species A, and A is therefore called an *absolute concentration robust* species. Such species are important in biological mechanisms because they provide a fixed response at equilibrium, regardless the environmental conditions, and therefore cannot be the cause of biological "switches". Structural sufficient conditions to recognize

such species are studied in Shinar & Feinberg (2010). In our example, let the fixed concentration of A at equilibrium be α, and let the sum of the initial concentrations of A and B be N. We have that for any positive initial condition satisfying $N > \alpha$, the limit of the solution of the ODE is $(\alpha, N - \alpha)$, where the first and the second entries denote the concentration of A and B, respectively. On the other hand, if we consider the stochastic mass-action system, we have that the process is absorbed with probability 1 by the state $(N, 0)$. The limit behaviour of the two models is therefore very different, especially if we imagine N as a large number. A more general statement is made in Anderson *et al.* (2014b), and it is proven that a large class of reaction networks share the observed discrepancy property. Such class is characterized by structural features of the reaction graph. The question left opened by Anderson *et al.* (2014b) is the magnitude of the time at which the divergence is revealed, and how different the behaviour of the two models is up to to a fixed finite time T. Such question is addressed in Paper III, where we prove that, up to any fixed time T, the stochastic and the deterministic dynamics are indeed very close. To this goal, we make use of averaging techniques studied in Kurtz (1992).

1.4 Thesis structure

In Chapter 2 the necessary definition from CRNT are given. In particular, the formal definition of the models introduced here are presented, as well as all the necessary background to discuss in more details our contribution in the international state-of-art of CRNT. Chapter 3 presents an accurate overview of the papers and their relation with other existing results. Some interesting perspectives and possible future works are discussed in Chapter 4, after which the bibliography with the references used in Chapters 1-4 can be found. Finally, the papers containing our contributions are collected, each one equipped with its own bibliography.

2

Mathematical background

In this chapter the necessary definitions and results in CRNT are introduced (for general references see Feinberg, 1987; Érdi & Tóth, 1989; Anderson & Kurtz, 2011, 2015). All the used notations and the needed definitions from CRNT are also repeated in the introductory sections of the papers collected here.

2.1 Notation

We let $\mathbb{R}$, $\mathbb{R}_0$ and $\mathbb{R}_+$ be the real, the non-negative real and the positive real numbers, respectively. Also let $\mathbb{N}$ be the natural numbers including 0.

For any real number $a \in \mathbb{R}$, $|a|$ denotes the absolute value of a. Moreover, for any vector $v \in \mathbb{R}^p$, we let v_i be the ith component of v, $\|v\|$ the Euclidean norm, $\|v\|_1$ its L^1-*norm* and $\|v\|_\infty$ its L^∞-*norm*, that is, $\|v\| = \sqrt{\sum_i v_i^2}$, $\|v\|_1 = \sum_i |v_i|$ and $\|v\|_\infty = \max_i |v_i|$. We denote by $[v]$ the vector of the floor functions of the entries of v; that is, $[v]_i = \lfloor v_i \rfloor$. For two vectors $v, w \in \mathbb{R}^p$, we write $v < w$ (resp. $v > w$) and $v \leq w$ (resp. $v \geq w$), if the inequality holds component-wise. Further, we define $\mathbb{1}_{\{v \leq w\}}$ to be one if $v \leq w$, and zero otherwise, and similarly for the other inequalities. If $v > 0$ then v is said to be positive. Moreover, $\langle v, w \rangle$ denotes the usual scalar product and $\operatorname{supp} v$ denotes the index set of the non-zero components of v. For example, if $v = (0, 1, 1)$ then $\operatorname{supp} v = \{2, 3\}$. Finally, we define

$$v^w = \prod_{i=1}^{p} v_i^{w_i}, \quad \text{and} \quad v! = \prod_{i=1}^{p} v_i!,$$

with the conventions that $0! = 1$ and $0^0 = 1$. Moreover, for any real number $N > 0$, we denote by $N^w v$ the vector satisfying

$$(N^w v)_i = N^{w_i} v_i.$$

Finally, for any set A we will indicate by $|A|$ its cardinality and for any $a, b \in \mathbb{R}$, $a \wedge b$ and $a \vee b$ will denote $\min\{a, b\}$ and $\max\{a, b\}$, respectively.

2.2 Graph theory and probability theory

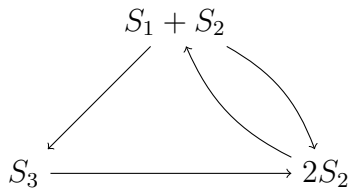
We assume basic knowledge on graphs. Specifically, we assume that it is known what a directed graph, a directed path and a directed cycle are. Familiarity with the concepts of connected and strongly connected components of a graph is useful, but not required.

In this thesis different topics from probability theory will be mentioned and used in connection to stochastic reaction systems (see 2.4.2). The theory we need is covered by Master courses in most universities, and we assume it is known. For convenience, in this section we summarize the main topics we will deal with. First, stochastic reaction systems are *continuous-time Markov chains*. We will then deal with the particular case of *Poisson processes*, and investigate the existence and the form of the *stationary distributions*. As a reference on Markov chain theory we propose Norris (1998); the definition as well as the meaning of the stationary distribution in our particular case are proposed in section 2.4.4. Finally, we will deal with the *weak convergence* of stochastic processes, which is intended to be with respect to the *Skorohod topology* unless otherwise stated. For a reference, we suggest consulting Ethier & Kurtz (2009).

2.3 Reaction networks

A reaction network is a triple $\mathcal{G} = (\mathcal{X}, \mathcal{C}, \mathcal{R})$, where $\mathcal{X}$ is a finite ordered set of symbols, referred to as *species*, and $\mathcal{C}$ is a finite ordered set of linear combinations of species on $\mathbb{N}$, referred to as *complexes*. Any species $S_i \in \mathcal{X}$ can be identified with the vector $e_i \in \mathbb{R}^{|\mathcal{X}|}$, whose ith entry is 1 and whose other entries are zero. Therefore, any complex $y \in \mathcal{C}$ will be identified with a vector in $\mathbb{R}^{|\mathcal{X}|}$ that is linear combination of the vectors e_i. Finally, $\mathcal{R}$ is a non-empty ordered subset of $\mathcal{C} \times \mathcal{C}$, whose elements are called *reactions*, such that for any $y \in \mathcal{C}$, $(y, y) \notin \mathcal{R}$. Following the usual notation, we will denote any element $(y_r, y_r') \in \mathcal{R}$ by $y_r \to y_r' \in \mathcal{R}$, in which case we then call y_r the *source complex* and y_r' the *product complex* of that reaction. It is possible that a complex $y \in \mathcal{C}$ is the source (product) complex of different reactions, and that it is both the source complex of one reaction and the product complex of another reaction. It is commonly required that every species $S \in \mathcal{X}$ appears in at least one complex, and that every complex $y \in \mathcal{C}$ appears as an element in at least one reaction. It is possible to associate a directed graph with $\mathcal{G}$, where the set of nodes is the set of complexes $\mathcal{C}$ and the arrows are given by the reactions $y_r \to y_r' \in \mathcal{R}$. Such graph is called *reaction graph* and uniquely determines the reaction network, such that reaction networks are usually introduced by means of their reaction graphs. As an example, consider the reaction network determined

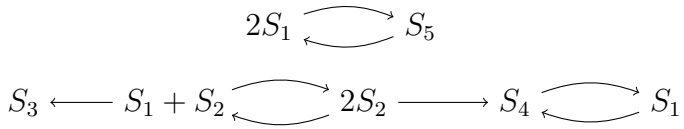
by



(2.1)

In this case, the set of species is $\mathcal{X} = \{S_1, S_2, S_3\}$, the set of complexes is $\mathcal{C} = \{S_1 + S_2, 2S_2, S_3\}$, the set of reactions is $\mathcal{R} = \{S_1 + S_2 \to 2S_2, S_1 + S_2 \to S_3, S_3 \to 2S_2, 2S_2 \to S_1 + S_2\}$, and each complex is both a source complex and a product complex of some reaction.

If the reaction graph of $\mathcal{G}$ is such that for any directed path from a complex y to another complex y' there exists a directed path from y' to y, then $\mathcal{G}$ is *weakly reversible*. For example, (2.1) is weakly reversible. The connected components of the graph determine a partition of the complexes into different *linkage classes*. For example, consider



(2.2)

The graph has two connected components, and the complexes are accordingly partitioned into the linkage classes $\{S_3, S_1 + S_2, 2S_2, S_4, S_1\}$ and $\{2S_1, S_5\}$. Formally, we can introduce the relation $\rightsquigarrow$ on the set of complexes, such that $y \rightsquigarrow y'$ if $y = y'$ or if there exists a directed path from y to y' in the reaction graph. Then, we define the equivalence relation $\sim$ by

$$y \sim y' \quad \text{if} \quad y \rightsquigarrow y' \quad \text{or} \quad y' \rightsquigarrow y.$$

The linkage classes are the equivalence classes of $\sim$. We can further define the equivalence relation $\simeq$ by

$$y \simeq y' \quad \text{if} \quad y \rightsquigarrow y' \quad \text{and} \quad y' \rightsquigarrow y.$$

The associated equivalence classes are termed *strong linkage classes*, and for the network (2.2) they are given by $\{S_3\}$, $\{S_1 + S_2, 2S_2\}$, $\{S_4, S_1\}$ and $\{2S_1, S_5\}$. Furthermore, a *terminal strong linkage class* is a strong linkage class L with no directed path out of it, that is such that for any $y \in L$

$$y \rightsquigarrow y' \Rightarrow y' \in L.$$

There are three terminal strong linkage classes in (2.2), that is $\{S_3\}$, $\{S_4, S_1\}$ and $\{2S_1, S_5\}$. A complex is *terminal* if it belongs to a terminal strong linkage class.

Finally, a reaction $y \to y' \in \mathcal{R}$ is *terminal* if any directed path that starts with $y \to y'$ is contained in a closed directed path. Equivalently, a terminal reaction could be defined as a reaction whose source complex is terminal. It follows from the definition of terminal complexes that the product complex of a terminal reaction is terminal as well. In (2.2), the terminal reactions are $S_1 \to S_4$, $S_4 \to S_1$, $2S_1 \to S_5$ and $S_5 \to 2S_1$. The subgraph determined by the terminal reactions is called the *terminal network* and in this case is given by

$$S_4 \rightleftarrows S_1 \qquad 2S_1 \rightleftarrows S_5$$

The terminal network is always weakly reversible, and an alternative definition of weak reversibility can be formulated in the following way: a reaction network $\mathcal{G}$ is weakly reversible if all its reactions are terminal. In such case, the terminal network coincides with $\mathcal{G}$.

For each reaction $y \to y'$, the vector $y' - y$, named *reaction vector*, corresponds to the net gain of molecules given by the occurrence of the reaction. In Papers I and III we denote the reaction vectors by the symbol ξ. The *stoichiometric subspace* of $\mathcal{G}$ is the linear subspace of $\mathbb{R}^{|\mathcal{X}|}$ given by

$$\mathcal{S} = \text{span}(y' - y | y \to y' \in \mathcal{R}).$$

For $v \in \mathbb{R}^{|\mathcal{X}|}$, the sets $(v + \mathcal{S}) \cap \mathbb{R}_0^{|\mathcal{X}|}$ are called the *stoichiometric compatibility classes* of $\mathcal{G}$. For the network in (2.1), $\mathcal{S} = \text{span}((-1, 1, 0), (0, 1, -1)) \subset \mathbb{R}^3$, which is 2-dimensional, while for (2.2) we have $\mathcal{S} = \mathbb{R}^5$. If there exists a positive vector orthogonal to $\mathcal{S}$, then the reaction network is called *conservative*. Equivalently, a reaction network is conservative if every stoichiometric compatibility class is bounded. For example, (2.1) is conservative because $(1, 1, 1)$ is orthogonal to the stoichiometric subspace, while (2.2) is not conservative because the unique stoichiometric compatibility class $\mathbb{R}_0^5$ is not bounded.

2.4 Dynamical systems

We will consider a reaction network $\mathcal{G}$ either as a deterministic dynamical system on the continuous space $\mathbb{R}_0^{|\mathcal{X}|}$, or as a stochastic dynamical system on the discrete space $\mathbb{N}^{|\mathcal{X}|}$.

2.4.1 Deterministic reaction systems

The deterministic model is normally used to describe biochemical machanisms with a lage abundance of chemical reactants, such that the concentration of each reactant

is considered, instead of the counts of the molecules present. We denote by $z(t)$ the non-negative real vector of $\mathbb{R}_0^{|\mathcal{X}|}$ whose entries are the concentrations of interest at time t, and we call every vector of $\mathbb{R}_0^{|\mathcal{X}|}$ a *state*. The evolution of $z(t)$ is modelled as the solution to the ODE

$$z(t) = \sum_{y \to y' \in \mathcal{R}} (y' - y) \int_0^t \lambda_{y \to y'}(z(s)) ds, \tag{2.3}$$

for some functions $\lambda_{y \to y'} \colon \mathbb{R}_0^{|\mathcal{X}|} \to \mathbb{R}_0$ and an initial condition $z(0) \in \mathbb{R}_0^{|\mathcal{X}|}$. We require that the functions $\lambda_{y \to y'}$ are continuously differentiable, and that $\lambda_{y \to y'}(z) > 0$ if and only if $\operatorname{supp} y \subseteq \operatorname{supp} z$. Such functions are called *rate functions*, they constitute a *deterministic kinetics* K for $\mathcal{G}$, and the pair $(\mathcal{G}, K)$ is called a *deterministic reaction system*. The rate functions express the velocity of the different reactions, as function of the concentrations of the different species.

It follows from (2.3) that $z(\cdot)$ is confined to the stoichiometric compatibility classes: for any $t \geq 0$

$$z(t) \in (z(0) + \mathcal{S}) \cap \mathbb{R}_0^{|\mathcal{X}|}.$$

In this sense, when existence and uniqueness of a steady state are investigated, the addressed question is whether a steady state exists and is unique within each stoichiometric compatibility class (Horn & Jackson, 1972; Feinberg, 1987, 1988). We discussed in Section 1.2 that the possibility of multiple (stable) steady states is also of interest, because it allows biological mechanisms to change their equilibrium configuration in response to different external stimuli. The question addressed when studying multistationarity is whether there exists a choice of kinetics (usually with certain restrictions) that implies the existence of more than a (stable) steady state on a single stoichiometric compatibility class (Feinberg, 1988; Markevich *et al.*, 2004; Thomson & Gunawardena, 2009; Feliu & Wiuf, 2013; Müller *et al.*, 2013; Joshi & Shiu, 2013; Feliu, 2015).
Note that in the case of conservative reaction networks the solution $z(\cdot)$ is bounded because it lies within a bounded stoichiometric compatibility class, regardless the initial condition $z(0)$.

If $\lambda_{y \to y'}(z) = \kappa_{y \to y'} z^y$ for all reactions for some positive constants $\kappa_{y \to y'}$, then the $\kappa_{y \to y'}$'s are referred to as *rate constants* and the modelling regime is referred to as *deterministic mass-action kinetics*. In this case, the pair $(\mathcal{G}, \kappa)$ is called a *deterministic mass-action system*, where $\kappa \in \mathbb{R}_+^{|\mathcal{R}|}$ is the vector of rate constants. As an example, the mass-action rate function for the reaction $A + B \to C$ has the form $\lambda_{A+B \to C}(z) = \kappa_{A+B \to C} z_A z_B$ for a positive constant $\kappa_{A+B \to C}$, where z_A and z_B denote the entries of z relative to the species A and B, respectively. The mass-action rate function for $2A \to 2B$ is quadratic in A and it is given by $\lambda_{2A \to 2B}(z) = \kappa_{2A \to 2B} z_A^2$, for a positive constant $\kappa_{2A \to 2B}$.

The mass-action kinetics assigns to each reaction a velocity that is proportional to the concentrations of the species of the source complex, and therefore describes well the dynamics of biochemical mechanism where the chemical components are well-mixed.

As a remark, we stress that it is generally not clear whether $z(t)$ exists for any non-negative t: if we consider the reaction network $2A \to 3A$ endowed with mass-action kinetics, if the initial condition is positive then $z(\cdot)$ tends to infinity in a finite time.

2.4.2 Stochastic reaction systems

The stochastic model is typically chosen when few molecules are present, so the focus is on the evolution of their counts. This is similar to a jump process, indeed reactions occur only in a discrete set of random time points and it is uncertain which one will take place next. Let $X(t)$ be the vector in $\mathbb{N}^{|\mathcal{X}|}$ whose entries are the species counts at time t. The evolution of $X(t)$ is usually modelled as a continuous-time Markov chain with state space $\mathbb{N}^{|\mathcal{X}|}$. At any state $x \in \mathbb{N}^{|\mathcal{X}|}$, the states that can be reached in one step are $x + y' - y$ for $y \to y' \in \mathcal{R}$, with transition rates $\lambda_{y\to y'}(x)$. The functions $\lambda_{y\to y'} \colon \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}_0$ are called *rate functions*, and we require that $\lambda_{y\to y'}(x) > 0$ if and only if $x \geq y$. This requirement essentially allows a reaction to take place whenever the molecules indicated by the source complex are present. For example, we want that the reaction $2A \to 2B$ may take place whenever at least two molecules of A are present, otherwise it cannot occur. A choice of such rate functions constitute a *stochastic kinetics* K for $\mathcal{G}$ and the pair $(\mathcal{G}, K)$ is called a *stochastic reaction system* (the latter definition is not universally accepted, and in Paper I we will simply talk about reaction networks with a kinetics).
If the reaction $y \to y'$ occurs at time t, then the new state is

$$X(t) = X(t-) + y' - y,$$

where $X(t-)$ denotes the previous state. Following the terminology utilized in Kurtz (1978), Ethier & Kurtz (2009), Anderson & Kurtz (2011) and Anderson & Kurtz (2015), we can write

$$X(t) = X(0) + \sum_{y\to y' \in \mathcal{R}} (y' - y) Y_{y\to y'} \left(\int_0^t \lambda_{y\to y'}(X(s)) ds \right), \tag{2.4}$$

with $Y_{y\to y'}$ independent and identically distributed unit-rate Poisson processes. The random variable $Y_{y\to y'} \left(\int_0^t \lambda_{y\to y'}(X(s)) ds \right)$ counts how many times the reaction $y \to y'$ has occurred up to time t. It is worth pointing out the similarity between the equations (2.3) and (2.4): they only differ for the Poisson processes $Y_{y\to y'}$.

It follows from (2.4) that, similarly to the deterministic case, $X(\cdot)$ is confined within the stoichiometric compatibility class: for any $t \geq 0$

$$X(t) \in (X(0) + \mathcal{S}) \cap \mathbb{R}_0^{|\mathcal{X}|}.$$

In fact, $X(t) \in (X(0)+\mathcal{S}) \cap \mathbb{N}^{|\mathcal{X}|}$, as $X(t)$ takes values in $\mathbb{N}^{|\mathcal{X}|}$. In particular, in case of conservative networks, the continuous-time Markov chain $X(\cdot)$ given the initial condition $X(0)$ has a finite number of states, and properties like the existence of a stationary distribution are automatically fulfilled.

If for any reaction $y \to y' \in \mathcal{R}$

$$\lambda_{y \to y'}(x) = \kappa_{y \to y'} \frac{x!}{(x-y)!} \mathbb{1}_{\{x \geq y\}},$$

then the constants $\kappa_{y \to y'}$ are known as *rate constants*, as in the deterministic case, and the modelling regime is referred to as *stochastic mass-action kinetics*. The pair $(\mathcal{G}, \kappa)$ is, in this case, called a *stochastic mass-action system*. The mass-action rate functions are proportional to the possible combinations of the present molecules that can give rise to the reaction, and correspond to the hypothesis of a well-stirred biochemical system. For example, the stochastic mass-action rate function for the reaction $A+B \to C$ has the form $\lambda_{A+B \to C}(x) = \kappa_{A+B \to C} x_A x_B$ for a positive constant $\kappa_{A+B \to C}$, while that for $2A \to 2B$ is given by $\lambda_{2A \to 2B}(x) = \kappa_{2A \to 2B} x_A (x_A - 1)$, for a positive constant $\kappa_{2A \to 2B}$. Note the difference with the deterministic mass-action rate functions.

Finally, fix a time $T > 0$. As for the deterministic case, it is not clear whether $X(\cdot \wedge T)$ is stochastically bounded, meaning that almost surely $\sup_{t \in [0,T]} \|X(t)\| < \infty$. As a matter of fact, the stochastic mass-action system given by $2A \to 3A$ has a positive probability to reach infinity in a finite time for any choice of rate constant, given that the initial condition is strictly positive.

2.4.3 Classical scaling

The stochastic and the deterministic models have something in common. The original purpose of the two models is actually to model the same biological phenomenon under different abundances of the involved chemical reactants: if few molecules are present, then a stochastic reaction system is considered, while if the molecules are in high abundance and the concentrations are taken into account rather than the counts, a deterministic reaction system is normally used. Since the underlying modelled phenomenon, which is described by the reaction network, is the same, we might expect some connection between the two models, or at least hope for it. The first proven relationship appears in Kurtz (1972) and is further generalized in Kurtz

(1978). It states that, under the hypothesis of mass-action kinetics, the deterministic reaction system is the *weak limit* of the stochastic reaction systems, if both the counts of the species and the volume of the modelled mechanism scale with a parameter N tending to infinity. Here we give an informal presentation of the result, in the case of mass-action kinetics.

Imagine that few molecules are reacting in a well-stirred closed compartment of unitary volume. We model the dynamics of the reacting molecules by means of a stochastic mass-action system with a specific reaction network and a certain choice of rate constants. By (2.4), the associated process $X(\cdot)$ is given by

$$X(t) = X(0) + \sum_{y \to y' \in \mathcal{R}} (y' - y) Y_{y \to y'} \left(\kappa_{y \to y'} \int_0^t \frac{X(s)!}{(X(s) - y)!} \mathbb{1}_{\{X(s) \geq y\}} ds \right)$$

Now assume that the volume of the compartment is increased to N, where N is a positive natural number. In a larger volume, it is harder for the floating molecules to find each other, and the rate constants should be therefore rescaled accordingly. Specifically, the probability that $n \geq 1$ particular molecules meet in a volume N in a specific time interval is equal to the probability that they meet in a volume 1, rescaled by N^{n-1}. Moreover, if there is a spontaneous production of some species, i.e. reactions with 0 as source complex, it seems legit to assume that in a larger volume a higher generation is observed, in particular that its rate is proportional to the volume. It follows that the rate constants for the new setting are given by

$$\kappa^N_{y \to y'} = N^{1 - \|y\|_1} \kappa^N_{y \to y'}.$$

Assume that the initial condition $X^N(0)$ also scale with N. In particular, assume that $N^{-1} X^N(0)$ converges weakly to a non-negative real constant $z(0)$, for N going to infinity. Let $X^N(\cdot)$ be the process of the stochastic reaction network with the new parameters, and consider the vector of concentrations at time t

$$\widehat{X}^N(t) = N^{-1} X^N(t).$$

The distribution of $\widehat{X}^N(\cdot)$ is given by

$$\widehat{X}^N(t) = N^{-1} X^N(0) + \sum_{y \to y' \in \mathcal{R}} (y' - y) N^{-1} Y_{y \to y'} \left(\kappa^N_{y \to y'} \int_0^t \frac{X^N(s)!}{(X^N(s) - y)!} \mathbb{1}_{\{X(s) \geq y\}} ds \right).$$

If the magnitude of $X^N(t)$ is that of N and N is large, we can approximate the latter by

$$\widehat{X}^N(t) \approx N^{-1} X^N(0) + \sum_{y \to y' \in \mathcal{R}} (y' - y) N^{-1} Y_{y \to y'} \left(N \kappa_{y \to y'} \int_0^t \left(\widehat{X}^N(s) \right)^y ds \right).$$

Inspired by the Law of Large Numbers for Poisson processes, we can intuitively approximate further to

$$\widehat{X}^N(t) \approx N^{-1}X^N(0) + \sum_{y\to y'\in\mathcal{R}} (y'-y)\kappa_{y\to y'} \int_0^t \left(\widehat{X}^N(s)\right)^y ds,$$

which resembles the solution (2.3) of the deterministic mass-action system, provided that $N^{-1}X^N(0)$ converges to $z(0)$. As a matter of fact, we have that for any $T > 0$, the rescaled process $\widehat{X}^N(\cdot \wedge T)$ converges to the ODE solution $z(\cdot)$ of the deterministic mass-action system, under the assumption that $z(t)$ exists for any t in the compact time interval $[0, T]$. This result, known as "classical scaling", is formally proved in Kurtz (1972) and Kurtz (1978). In Kurtz (1978) a stronger result is actually shown, since no mass-action kinetics is assumed. Under some very general technical assumptions, the weak convergence holds for a family of stochastic reaction systems whose stochastic rate functions $\lambda^N_{y\to y'}(\cdot)$ are such that for any $v \in \mathbb{R}_0^{|\mathcal{X}|}$

$$\lim_{N\to\infty} N^{-1}\lambda^N_{y\to y'}(\lfloor Nv \rfloor) = \lambda_{y\to y'}(v),$$

for certain locally Lipschitz functions $\lambda_{y\to y'}(\cdot)$. These functions will constitute the kinetics for the deterministic reaction system whose solution $z(\cdot)$ is the weak limit of the rescaled stochastic processes $\widehat{X}^N(\cdot)$, up to a finite time T. It is worth noting that the result is about the transient behaviour of the stochastic reaction system, and the stationary behaviour may differ gratly from that of $z(\cdot)$ (see Section 3.3 for an example).

2.4.4 Stationary distribution

Here we recall some notion from Markov chain theory and apply it to the case of stochastic reaction networks. First, we introduce some terminology and some state characterisation, then we proceed with the definition of stationary distribution.

Definition 1. *Let $\mathcal{G} = (\mathcal{X}, \mathcal{C}, \mathcal{R})$ be a reaction network.*

a) A reaction $y \to y' \in \mathcal{R}$ is active *on $x \in \mathbb{N}^{|\mathcal{X}|}$ if $x \geq y$.*

b) A state $u \in \mathbb{N}^{|\mathcal{X}|}$ is accessible *from a state $x \in \mathbb{N}^{|\mathcal{X}|}$ if there is a sequence of $q \geq 0$ reactions $(y_j \to y'_j)_{j=1,\dots,q}$ such that*

(i) $u = x + \sum_{j=1}^{q}(y'_j - y_j)$,

(ii) $y_j \to y'_j$ is active on $x + \sum_{j=1}^{h-1}(y'_j - y_j)$ for all $1 < h \leq q$.

c) A non-empty set $\Gamma \subseteq \mathbb{N}^{|\mathcal{X}|}$ *is an* irreducible component *of* $\mathcal{G}$ *if for all* $x \in \Gamma$ *and all* $u \in \mathbb{N}^{|\mathcal{X}|}$, *u is accessible from x if and only if* $u \in \Gamma$.

d) $\mathcal{G}$ *is* essential *if the state space is a union of irreducible components.* $\mathcal{G}$ *is* almost essential *if the state space is a union of irreducible components except for a finite number of states.*

'Accessible', 'irreducible' and 'essential' are standard terms in Markov chain theory. A reaction network is essential if and only if every state of the associated Markov chain is 'essential' (Gusak *et al.*, 2010). In Paulevé *et al.* (2014) essential networks are termed 'recurrent', but here we do not follow this terminology as it may lead to confusion. Indeed, the states of an essential network are not necessarily recurrent in the usual sense of Markov chain theory, meaning that they will be visited again with probability 1 if they are chosen as initial condition. This is shown by the following stochastic mass-action system:

$$2S_1 \xrightarrow{\kappa_1} 3S_1 \qquad\qquad S_1 \underset{\kappa_3}{\overset{\kappa_2}{\rightleftarrows}} 0 \tag{2.5}$$

We wrote the rate constants next to the arrow of the corresponding reaction. This is usually done in CRNT and we will use this notation further on. The associated Markov chain is a 'birth-death process' and due to some classical results no state is recurrent in this case (Norris, 1998). Neverthless, the underlying reaction network is essential, because every state of $\mathbb{N}$ is accessible from any other state.

A weakly reversible reaction network is essential (Paulevé *et al.*, 2014), and further conditions for being essential can be found in Paulevé *et al.* (2014). Any irreducible component is contained in some stoichiometric compatibility class, and a stoichiometric compatibility class may contain several irreducible components (Fig. 1*B*)

The stationary distribution π_Γ on an irreducible component Γ is unique, if it exists. It is characterised by the *master equation* (Anderson & Kurtz, 2011):

$$\sum_{y \to y' \in \mathcal{R}} \pi_\Gamma(x + y - y')\lambda_{y \to y'}(x + y - y') = \pi_\Gamma(x) \sum_{y \to y' \in \mathcal{R}} \lambda_{y \to y'}(x), \tag{2.6}$$

for all $x \in \Gamma$. If $X(t_0)$ follows the law of π_Γ at time t_0, then the distribution of $X(t)$ is π_Γ for all future times $t \geq t_0$. In this sense, the stationary distribution describes a state of equilibrium of the system. Moreover, if π_Γ exists, then

$$\lim_{t \to \infty} P(X(t) \in A) = \pi_\Gamma(A) \quad \text{for any } A \subseteq \Gamma, \tag{2.7}$$

provided that $X(0) \in \Gamma$ with probability one. In this sense, stationary distributions can be regarded as the stochastic counterpart of the steady states of deterministic reaction systems: if $z(t_0)$ is a steady state, then $z(t) = z(t_0)$ for all future times $t \geq t_0$. Moreover, steady states are often limit points of ODE solutions. Therefore, if our aim is to compare the different dynamics of stochastic and deterministic reaction systems, it might be a good idea to study the connections between stationary distributions and steady states. This is done in Anderson *et al.* (2010), Anderson *et al.* (2014a), Anderson *et al.* (2014b) and Joshi (2015) as well as in Paper II. It is worth underlying that in general it is not easy to solve (2.6) and calculationg the stationary distribution can be hard, so additional information on its form given by the steady states of the deterministic model may be useful. Remember that the classical scaling introduced in section 2.4.3 does not give any information on the relationship between the limit behaviours of the stochastic and the deterministic reaction systems, as the convergence is up to a fixed time T.

2.5 Deficiency theory

Deficiency theory was introduced in 1972 by Horn & Jackson (1972), Feinberg (1972) and Horn (1972) as a set of results concerning deterministic reaction networks. The focus is connecting structural properties of the reaction graph with dynamical features of the reaction systems, primarily under the hypothesis of mass-action kinetics. Specifically, the existence and uniqueness of asymptotically stable steady states within each stoichiometric compatibility class is investigated for systems whose network is *deficiency zero* and weakly reversible. In (Horn & Jackson, 1972, Lemma 4C), the existence of *globally* asymptotically stable steady states within each stoichiometric compatibility class was stated, but the argument contained a flow, which was identified in Horn (1974). The question on whether the asymptotically stable steady states of deficiency zero weakly reversible mass-action systems are also globally asymptotically stable gave rise to a conjecture that survived for more than 40 years and was recently positively solved by Craciun (2015).

Further studies concern higher deficiency networks (Feinberg, 1987, 1988; Boros, 2013), but they will not be presented in this thesis.

2.5.1 Complex balanced equilibria

The first milestone of deficiency theory is perhaps the concept of *complex balanced systems*, introduced in Horn & Jackson (1972) as a generalization of the *detailed balanced systems*, not treated in this thesis. The concept of complex balancing itself is not related to structural properties of the network and depends on the kinetics,

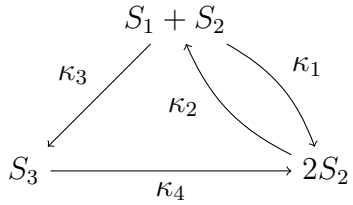
but important connections with weakly reversibility and deficiency have been proved. We begin with the definition of complex balanced systems and complex balanced equilibria.

Definition 2. *A deterministic reaction system* $(\mathcal{G}, K)$ *is said to be* complex balanced *if there exists a* complex balanced equilibrium*, that is, a positive equilibrium point* $c \in \mathbb{R}_+^{|\mathcal{X}|}$ *for the ODE* (2.3)*, such that*

$$\sum_{y' \in \mathcal{C}} \lambda_{y \to y'}(c) = \sum_{y' \in \mathcal{C}} \lambda_{y' \to y}(c) \quad \textit{for all } y \in \mathcal{C}., \tag{2.8}$$

with the convention that $\lambda_{y \to y'}(\cdot) = 0$ *if* $y \to y' \notin \mathcal{R}$*.*

The name 'complex balanced' refers to the fact that the flow, at equilibrium, entering into the complex y equals the flow exiting from the complex. As an example, consider (2.1) endowed with deterministic mass-action kinetics:



The associated deterministic mass-action system is complex balanced and

$$c = \left(\kappa_2, \kappa_1 + 2\kappa_3, \frac{\kappa_2 \kappa_3}{\kappa_4}(\kappa_1 + 2\kappa_3) \right)$$

is a complex balanced equilibrium.

The concept of complex balanced equilibria has been proved fruitful especially for the deterministic mass-action systems. Under the hypotheses of mass-action kinetics, (2.8) becomes

$$\sum_{y' \in \mathcal{C}} \kappa_{y \to y'} c^y = \sum_{y' \in \mathcal{C}} \kappa_{y' \to y} c^{y'} \quad \text{for all } y \in \mathcal{C}, \tag{2.9}$$

with the convention that $k_{y \to y'} = 0$ if $y \to y' \notin \mathcal{R}$.

We extend Definition 2 to the stochastic mass-action systems, by saying that a stochastic mass-action system $(\mathcal{G}, \kappa)$ is complex balanced if the deterministic mass-action system $(\mathcal{G}, \kappa)$ is complex balanced. We might therefore refer to complex balanced mass-action systems without specifying whether they are stochastically or deterministically modelled.

The next theorem is a classical result (Horn & Jackson, 1972), which provides the backbone for the further characterisation:

Theorem 1. *If a deterministic reaction system $(\mathcal{G}, K)$ is complex balanced, then $\mathcal{G}$ is weakly reversible. Moreover, if K is mass-action kinetics, all positive equilibria are complex balanced, that is, fulfil* (2.9)*. Moreover, there exists exactly one complex balanced equilibrium in each stoichiometric compatibility class, which is locally asymptotically stable.*

We have already mentioned that in the original version of the theorem, it is actually stated that complex balanced equilibria are *globally* asymptotically stable, but the argument proving this stronger property contained an error which was spotted and discussed in Horn (1974). However, the result as stated in Theorem 1 holds and is a classical result (Horn, 1974; Feinberg, 1979, 1987).

2.5.2 Deficiency zero

The *deficiency* plays an important role in the study of complex balanced systems. The deficiency of $\mathcal{G}$ is defined as

$$\delta = |\mathcal{C}| - \ell - s,$$

where ℓ is the number of the linkage classes of $\mathcal{G}$ and s is the dimension of the stoichiometric subspace $\mathcal{S}$ (Feinberg, 1972; Horn, 1972; Feinberg, 1987). For example, both the reaction networks in (2.1) and (2.2) have deficiency 0: in (2.1) we have $\delta = 3-1-2 = 0$, while in (2.2) we have $\delta = 7-2-5 = 0$. Oppositely, (2.5) has deficiency $\delta = 4 - 2 - 1 = 1$. By means of the above definition it is fairly easy to determine the deficiency of a network, even for a big one, but its geometrical interpretation stays hidden. Roughly speaking, the deficiency counts how many degrees of linear dependence between the reaction vectors are determined by the particular choice of complexes, and are not intrinsic in the structure of the graph itself. For example, in (2.1) the reaction vectors of $S_1 + S_2 \to 2S_2$ and $2S_2 \to S_1 + S_2$ are dependent, but this dependence is implied by the particular structure of the network, in the sense that if $S_1 + S_2$ and S_2 are substituted by any other choice of complexes y and y' in the corresponding nodes of the graph, the reaction vectors of $y \to y'$ and $y' \to y$ would still be linearly dependent. In other words, the dependence in this case comes from the fact the the reaction is reversible, which is a structural property of the reaction graph. On the other hand, if we consider the reaction network

$$S_1 + S_2 \longrightarrow 2S_2 \qquad\qquad S_2 \longrightarrow S_1 \tag{2.10}$$

the linear dependence of the reaction vectors of $S_1 + S_2 \to 2S_2$ and $S_2 \to S_1$ comes from the particular choice of complexes. Indeed, if we keep the same graph and we change S_2 with another complex, say $2S_1$, then the two reaction vectors of the

network are not dependent anymore. In (2.10), a linear dependence between the reaction vectors exists, which is not determined by the structure of the graph. In this case, the deficiency of the network is 1. We explore this concept more formally: let $\{e_y\}_{y\in\mathcal{C}}$ be a basis of $\mathbb{R}^{|\mathcal{C}|}$. Further, define

$$d_{y\to y'} = e_{y'} - e_y \quad \text{and} \quad \xi_{y\to y'} = y' - y \tag{2.11}$$

for $y \to y' \in \mathcal{R}$. Let $D = \operatorname{span}(d_{y\to y'} | y \to y' \in \mathcal{R})$. Then $\dim D = m - \ell$ (Feinberg, 1972). We have that the space D is linearly isomorphic to the stoichiometric subspace $\mathcal{S}$ if and only if $\delta = 0$. Specifically, consider the homomorphism

$$\begin{array}{rccc} \varphi\colon & \mathbb{R}^{|\mathcal{C}|} & \to & \mathbb{R}^{|\mathcal{X}|} \\ & e_y & \mapsto & y. \end{array} \tag{2.12}$$

For $y \to y' \in \mathcal{R}$, we have $\varphi(d_{y\to y'}) = \xi_{y\to y'}$ and $\varphi_{|D}\colon D \to \mathcal{S}$ is thus a surjective homomorfism. Therefore,

$$\delta = \dim D - s = \dim \operatorname{Ker} \varphi_{|D}, \tag{2.13}$$

which implies that $\varphi_{|D}$ is an isomorphism if and only if $\delta = 0$. It further follows that the deficiency is a non-negative number.

We next state a classical results which elucidate the connection between complex balanced systems and deficiency zero systems (Horn, 1972; Feinberg, 1972).

Theorem 2. *The mass-action system $(\mathcal{G}, \kappa)$ is complex balanced for any choice of κ if and only if $\mathcal{G}$ is weakly reversible and its deficiency is zero.*

Another classical result is known by the name of 'Deficiency Zero Theorem' (Feinberg, 1979, 1987). To state it, we first define *cyclic composition trajectory* of an ODE as a closed directed path composed by concatenating finitely many trajectories of the ODE. A cyclic composition trajectory can be for example a steady state, a periodic orbit, a homoclinic cycle or a heteroclinic cycle. A cyclic composition trajectory is *non-trivial* if it is not a single steady state.

Theorem 3 (Deficiency zero theorem)**.** *Consider a deterministic reaction system $(\mathcal{G}, K)$ for which the deficiency is zero. Then the following statements hold:*

i) if $\mathcal{G}$ is not weakly reversible, then the associated ODE admits no positive steady state;

ii) if $\mathcal{G}$ is not weakly reversible, then the associated ODE admits no cyclic composition trajectory containing a positive state (i.e. a positive vector of $\mathbb{R}^{|\mathcal{X}|}$);

iii) if $\mathcal{G}$ is weakly reversible and K is mass-action kinetics, then there exists within each stoichiometric compatibility class a unique positive equilibrium, which is asymptotically stable. Moreover, there cannot exists a non-trivial cyclic composition trajectory in $\mathbb{R}_+^{|\mathcal{X}|}$.

In case of non-weakly reversibility, something more precise can be stated (Feinberg, 1987):

Theorem 4. *Consider a deterministic reaction system $(\mathcal{G}, K)$, and assume that the deficiency of $\mathcal{G}$ is zero. If $x \in \mathbb{R}_0^{|\mathcal{X}|}$ is a steady state and $y \in \mathcal{C}$, then* $\operatorname{supp} y \subseteq \operatorname{supp} x$ *only if y is terminal.*

In Paper II we prove a slightly expanded version of Theorem 4 (See Paper II, Theorem 7).

Theorem 3 can be applied to deterministic mass-action systems built on (2.1) and (2.2) to exclude any 'exotic' behaviour of the associate solutions $z(\cdot)$. In particular, any mass-action system built on (2.1) has exactly one locally asymptotically stable steady state within each stoichiometric compatibility class, and all cyclic composition trajectories in $\mathbb{R}_+^{|\mathcal{X}|}$ are a single steady state. Moreover, by Theorems 2 and 1 each steady state is complex balanced. Moreover, from the Deficiency Zero Theorem it follows that any deterministic reaction system built on (2.2) admits no cyclic composition trajectory containing a positive state, regardless the chosen kinetics. Furthermore, by Theorem 4 the second entry of every steady state is null, since $2S_2$ is a non-terminal complex in (2.2).
It is worth noting the amount of information we are able to gain on the dynamics of deterministic reaction systems by only looking at simple structural properties of the reaction network, such weakly reversibility and deficiency. Unfortunately things are not so pleasant for higher deficiency models, thought some important result has been proven for deficiency one reaction networks (Feinberg, 1987, 1988; Boros, 2013).

3

Overview of the results

In this chapther a detailed overview of the results in the collected papers and their relations with the existing theory is given. Some of the results that are proven in the papers are stated here, and a part of them is stated in a reduced form. No attempt is made to maintain the numeration that appears in the papers.

3.1 Overview of Paper I

The classical scaling described in Section 2.4.3 is the first attempt to connect the stochastic and the deterministic models. Moreover, it provides an approximation for the dynamics of the stochastic model by means of the ODE solution of deterministic counterpart when the counts of the species are high, up to a finite time T. Such an approximation is extremely useful because a high number of molecules prevents an efficient computational analysis of the stochastic system, since the computational time of the simulations is drastically increased. A natural question is whether a similar approximation can be performed in a *multiscale* setting, i.e. when the abundances of the different species strongly differ. This may happen in a model of a living cell for example, where it is not uncommon to see chemical reactants scaling with different orders of magnitude. Which kind of simplification can be performed in this case? The question has been addressed in Ball *et al.* (2006), Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013).
Formally, a multiscale setting can be defined by letting the counts of the different species in the setup of Section 2.4.3 to scale with different non-negative powers of N. Namely, consider a vector $\alpha \in \mathbb{R}_0^{|\mathcal{X}|}$ of non-negative powers, such that $N^{-\alpha}X^N(0)$ converges weakly to a non-negative real random variable $Z(0)$: the vector α expresses the magnitudes of the different chemical species. Moreover, we let the rate functions scale with powers of N as well: assume that there exists a non-negative vector $\beta \in \mathbb{R}_0^{|\mathcal{R}|}$, which we index for simplicity by the reactions in $\mathcal{R}$, such that, for any $y \to y' \in \mathcal{R}$,

$$N^{-\beta_{y\to y'}}\lambda^N_{y\to y'}(\lfloor N^\alpha x \rfloor) \xrightarrow[N\to\infty]{} \lambda_{y\to y'}(x) \tag{3.1}$$

uniformly on compact sets of $\mathbb{R}_0^{|\mathcal{X}|}$, where the function $\lambda_{y\to y'}\colon \mathbb{R}_0^{|\mathcal{X}|} \to \mathbb{R}_0$ is a locally Lipschitz function. For mass-action kinetics, if the constants $\kappa^N_{y\to y'}$ are of the form $N^{\eta_{y\to y'}}\kappa_{y\to y'}$ with $\kappa_{y\to y'} > 0$ and $\eta_{y\to y'} \in \mathbb{R}$, then

$$\lambda^N_{y\to y'}(N^\alpha x) = N^{\eta_{y\to y'}}\kappa_{y\to y'}\frac{(N^\alpha x)!}{(N^\alpha x - y)!}\mathbb{1}_{\{N^\alpha x \geq y\}}\,.$$

This means that the right scaling for the rate function $\lambda^N_{y\to y'}$ is

$$\beta_{y\to y'} = \eta_{y\to y'} + \langle \alpha, y\rangle\,.$$

Indeed, with this choice we have (3.1) uniformly on compact sets, where

$$\lambda_{y\to y'}(x) = \kappa_{y\to y'}\prod_{\substack{i\in \operatorname{supp} y\\ \alpha_i = 0}}\frac{x_i!}{(x_i-y_i)!}\mathbb{1}_{\{x_i\geq y_i\}}\prod_{\substack{i\in\operatorname{supp} y\\ \alpha_i>0}} x_i^{y_i}\,.$$

In order to state the convergence results, we need the following conditions:

Assumption 1. *Let T be a positive real number. We assume that:*

i) the magnitude of the function rates is bounded by the availability of the species whose amount is changed by the reaction, that is

$$\max_{y\to y'\in\mathcal{R}\colon \xi_{y\to y',i}\neq 0}\beta_{y\to y'} \leq \alpha_i\,, \tag{3.2}$$

where $\xi_{y\to y'}$ is as in (2.11), and $\xi_{y\to y',i}$ denotes its ith component;

ii) the rescaled process is uniformly stochastic bounded, meaning that for any $\nu > 0$, there exists a constant Υ_ν such that

$$\limsup_{N\to\infty} P\left(\sup_{[0,T]}\widehat{X}^N_t(S) > \Upsilon_\nu\right) < \nu\,;$$

iii) if we let

$$\xi^*_{y\to y'} = \lim_{N\to\infty} N^{\beta_{y\to y'}}N^{-\alpha}\xi_{y\to y'},$$

then the process

$$Z^*(t) = Z(0) + \sum_{\substack{y\to y'\in\mathcal{R}\\ \beta_{y\to y'}=0}}\xi^*_{y\to y'}Y_{y\to y'}\left(\int_0^t \lambda_{y\to y'}(Z^*(s))ds\right) + \sum_{\substack{y\to y'\in\mathcal{R}\\ \beta_{y\to y'}>0}}\xi^*_{y\to y'}\int_0^t\lambda_{y\to y'}(Z^*(s))ds \tag{3.3}$$

is well-defined up to T. In particular, its distribution is uniquely defined and $Z^*(\cdot)$ is bounded with probability 1 up to time T.

In our setting, the mass produced or consumed by a reaction $y \to y'$ is of the order of $N^{\beta_{y\to y'}}$ and the counts of a species S_i are of order N^{α_i}. Therefore, in Assumption 1(i) we allow for the rescaled amounts of species to change in a time scale that is equal to or larger than the one considered, meaning in the latter case that we do not observe any significant variation of the rescaled counts of the species in a finite time. Following the terminology of Pfaffelhuber & Popovic (2013), an equality in (3.2) defines a *single-scale* system: the term "single-scale" refers to the fact that all the entries of $N^{-\alpha}X^N(\cdot)$ evolve on the same time scale. Different time scales can be considered by simply scaling the rate functions by a common factor, typically a power of N: if for example the rate functions are all divided by N, then the dynamics on the time scale of $t' = Nt$ are considered. In this case, the entries of the vector β are all decreased by 1. The time scale of interest is explicitly stated in Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013) by means of a real number γ (the time scale will be then given by $N^\gamma t$), however it is ignored in this manuscript since the dynamics on any time scale can be considered by simply rescaling the rate functions. It is worth noting that some restrictions are imposed by Assumption 1(i), which prevents evolutions of the rescled amounts of species that happen in a faster time scale than the one observed. We have the following result (Kang & Kurtz, 2013; Pfaffelhuber & Popovic, 2013):

Theorem 5. *Assume that Assumption 1 holds. Then, the stopped process* $N^{-\alpha}X^N(\cdot \wedge T)$ *converges weakly to* $Z^*(\cdot \wedge T)$.

In the limit process $Z^*(\cdot)$, each reaction $y \to y'$ is associated with a vector $\xi^*_{y\to y'}$ which has a smaller support than $\xi_{y\to y'}$, due to Assumption 1(i). In particular, the reaction $y \to y'$ can be eliminated from the asymptotical approximation of the model if $\xi^*_{y\to y'} = 0$.

Techniques introduced in Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013) allow to handle systems with a set of species that does not satisfy (3.2). For example, assume that the set of species is partitioned into a set of *slow species* $\mathcal{X}_s$, for which (3.2) holds, and a set of *fast species* $\mathcal{X}_f$, for which

$$\max_{y\to y'\in\mathcal{R}\,:\,\xi_{y\to y',i}\neq 0} \beta_{y\to y'} = \alpha_i + \theta\,,$$

foa a certain common $\theta > 0$. Such a system is referred to as a *two-scale system* in Pfaffelhuber & Popovic (2013). The idea to simplify the system is that the fast species, whose changes happen at a faster time scale, reach a stationary distribution before that the slow species can significantly change their concentration. Let p_s and p_f be the canonical projections of $\mathbb{R}^{|\mathcal{X}|}$ onto the spaces of the slow and the fast species, respectively. For the sake of simplicity, we introduce the notation

$$\lambda_{y\to y'}(p_s(x), p_f(x)) = \lambda_{y\to y'}(x).$$

We define

$$\xi^{*,f}_{y\to y'} = \lim_{N\to\infty} N^{\beta_{y\to y'}} N^{-\alpha-\theta} p_f(\xi_{y\to y'}),$$

and assume that the process

$$Z^v_f(t) = p_f(Z(0)) + \sum_{\substack{y\to y'\in\mathcal{R}\\ \beta_{y\to y'}=0}} \xi^*_{y\to y'} Y_{y\to y'}\left(\int_0^t \lambda_{y\to y'}\left(v, Z^v_f(s)\right) ds\right) + \sum_{\substack{y\to y'\in\mathcal{R}\\ \beta_{y\to y'}>0}} \xi^*_{y\to y'} \int_0^t \lambda_{y\to y'}\left(v, Z^v_f(s)\right) ds$$

has a unique stationary distribution π^v for any $v \in \mathbb{R}_0^{|\mathcal{X}_s|}$. Then, assuming that all the integrals in the next definition are finite and under some other technical conditions, the stopped process $p_s(X^N(\cdot \wedge T))$ converges weakly to the process defined by

$$Z_s(t\wedge T) = p_s(Z(0)) + \sum_{\substack{y\to y'\in\mathcal{R}\\ \beta_{y\to y'}=0}} p_s(\xi^*_{y\to y'}) Y_{y\to y'}\left(\int_0^{t\wedge T}\int_{\mathbb{R}^{|\mathcal{X}_f|}} \lambda_{y\to y'}\left(Z_s(s), w\right) d\pi^{Z_s(s)}(w) ds\right) + \sum_{\substack{y\to y'\in\mathcal{R}\\ \beta_{y\to y'}>0}} p_s(\xi^*_{y\to y'}) \int_0^t \int_{\mathbb{R}^{|\mathcal{X}_f|}} \lambda_{y\to y'}\left(Z_s(s), w\right) d\pi^{Z_s(s)}(w) ds.$$

By using the same idea, we can approximate three-scale systems, four-scale systems and so on. See Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013) for further details.

Consider now the mass-action system given by

$$S_1 + S_2 \underset{\kappa_1 N^{\eta_1}}{\overset{\kappa_0}{\rightleftarrows}} S_3 \xrightarrow{\kappa_2 N^{\eta_2}} S_1 + S_4 \tag{3.4}$$

where κ_0, κ_1, κ_2, η_1 and η_1 are positive real numbers, while N is a large natural number. The system is a Michaelis-Menten mechanism where the substrate S_2 is transformed into the protein S_4 by means of the enzyme S_2 and proceeding through the formation of an intermediate species, which is fast degraded. If we let $\alpha_1 = \alpha_3 = 0$ and $\alpha_2 = \alpha_4 = 1$, then (3.2) does not hold for S_1 and S_2. If we assume $\eta_2 > 1$, then (3.2) does not hold for S_4 either. Therefore, by using the techniques introduced in Kang & Kurtz (2013); Pfaffelhuber & Popovic (2013) we can only approximate the evolution of the species S_2, and we lose track of all the other species. On the

other hand, in this case it is clear what happens for large values of N: whenever a molecule of S_3 is produced, it is fast degraded in such a way that most of the time no molecules of S_3 will be present, and with high probability at most one molecule of S_3 can be observed at any time. This gives an indication on the evolution of S_1, whose count is more or less constant, and we can also imagine to approximate the dynamics of S_4 by contracting the reaction paths in (3.4) to obtain

$$S_1 + S_2 \longrightarrow S_1 + S_4$$

with some kinetics. Paper I focuses on this type of reduction, which can not be performed in the settings of Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013) (see also Kang & Kurtz, 2013, Section 6.5). The reduction is motivated by the fact that intermediate species appear often in biochemical reaction networks (for example in the enzymatic processes), frequently they are not detectable and fast consumed. Indeed, modelers in biochemistry sometimes ignore intermediate species, but so far no formal result supporting this choice was available, or stating which properties are preserved under the simplification. We study different kind of convergence of the original process to the simplified one, both in the case where some boundedness properties hold and in the case where the simplified process has a weak limit (up to a certain fixed time T). Specifically, weak convergence holds if $\alpha > 0$, but not in general (see Theorem 4.4, Theorem 4.6 and Example 5.3 in Paper I). However, in general a uniform punctual weak convergence holds, as well as a weak convergence of the integrals of the two processes, in a sense that is made clear by Theorems 4.2 and 4.6.

Some previous work on intermediate species has been carried on by Feliu & Wiuf (2013), though the study was focused on steady states of deterministic reaction networks. Specifically, in the paper a system with intermediate species and a reduced system with no intermediate species are considered. The reactions and the kinetics of the reduced system are deduced by the full system following certain rules, and the main result states that the maximum number of steady states on a same stoichiometric compatibility class for the full system is greater than or equal to that of the reduced system. When inquiring for the possibility of multistationarity in a reaction system with intermediates species, this results is useful because allows to reduce the study to a simplified model. Although the areas of interest are different, we are able to prove in Paper I that the procedure we introduce to reduce a system with intermediates coincides with that utilized in Feliu & Wiuf (2013), implying that the same reduced model is able to provide information in different settings.

In Appendix IA, it is further shown that in the deterministic setting, intermediate species can be eliminated from the reaction system following the same procedure described in Paper I. The reduced system obtained is capable of uniformly approximate the trajectories of the original model, up to finite time T. Since the reduced

model coincides with that introduced by Feliu & Wiuf (2013), we have that the same reduced model can approximate both trajectories and steady states of the full system with intermediate species.

3.2 Overview of Paper II

In Section 2.5, we have seen a collection of results for deterministic reaction systems. These results link the deficiency of a network with dynamical properties of the system. No such theory have been developed for stochastic reaction systems, and the concept of *complex balanced equilibrium* does not have a stochastic counterpart. In Paper II we build such a theory and we introduce the concept of *complex balanced stationary distribution*. The theory we develop focuses on links between structural properties of a reaction network and features of the stationary distribution of an associated stochastic reaction system. The goal is indeed to state results for the stochastic systems that are parallel to those in Section 2.5, and in this process it seems natural to consider the stationary distributions as the "translation" in the stochastic setting of the concept of steady states, due to the similarities discussed in Section 2.4.4. The definition of a complex balanced stationary distribution closely relates to that of a complex balanced equilibrium:

Definition 3. *Let* $(\mathcal{G}, K)$ *be a stochastic reaction system. A stationary distribution on an irreducible component* Γ *is said to be* complex balanced *if*

$$\sum_{y \in \mathcal{C}_\Gamma} \pi_\Gamma(x - y' + y)\lambda_{y \to y'}(x - y' + y) = \sum_{y \in \mathcal{C}_\Gamma} \pi_\Gamma(x)\lambda_{y' \to y}(x) \quad \forall y' \in \mathcal{C}_\Gamma, x \in \Gamma. \quad (3.5)$$

There is something missing in this definition, and it is the positivity that is required for complex balanced equilibria (see Definition 2). We then call a irreducible component *positive* if any reaction in $\mathcal{R}$ is active on at least one state of Γ. We then give the following definition:

Definition 4. *A stochastic reaction system* $(\mathcal{G}, \kappa)$ *is said to be* stochastically complex balanced *if there exists a complex balanced stationary distribution on a positive irreducible component.*

The link between complex balanced equilibria and the newly introduced complex balanced stationary distribution is strong, and we can prove that a deterministic mass-action system is complex balanced if and only if, when stochastically modelled, it is stochastically complex balanced (see Paper II, Corollary 19). Moreover, we build a theory that is parallel to that presented in Section 2.5. We briefly present it here:

Theorem 6. *If a stochastic reaction system $(\mathcal{G}, K)$ is stochastically complex balanced then $\mathcal{G}$ is weakly reversible. If K is mass-action kinetics, then on every irreducible component Γ there exists a unique stationary distribution π_Γ, which is a complex balanced stationary distribution.*

Theorem 6 (see Paper II, Corollary 19) might be considered a stochastic version of Theorem 1, especially if (2.7) is taken to be equivalent to the concept of "asymptotic stability" for a deterministic equilibrium.

Theorem 7. *The mass-action system $(\mathcal{G}, \kappa)$ is stochastically complex balanced for any choice of κ if and only if $\mathcal{G}$ is weakly reversible and its deficiency is zero.*

Theorem 8. *Consider a stochastic reaction system $(\mathcal{G}, K)$, and assume that the deficiency of $\mathcal{G}$ is zero. Then the following statements hold:*

i) if $\mathcal{G}$ is not weakly reversible, then there exist no positive irreducible components;

ii) if $\mathcal{G}$ is weakly reversible, then $\mathcal{G}$ is essential, and if K is mass-action kinetics there exists a unique stationary distribution on every irreducible component.

Theorem 9. *Consider a stochastic reaction system $(\mathcal{G}, K)$, and assume the deficiency of $\mathcal{G}$ is zero. Let x be a state in an irreducible component Γ and let $y \in \mathcal{C}$. Then, $y \leq x$ only if y is terminal.*

Theorems 6-9 are the stochastic correspondences of Theorems 1-4. In Paper II the theory is stated in a slightly different and richer way, and in particular stronger versions of Theorems 4 and 9 are proved.

Complex balanced stationary distributions for stochastic mass-action systems are always rescaled product-form Poisson distributions (see Paper II, Theorem 18), namely the unique stationary distribution on a irreducible component Γ has the form

$$\pi_\Gamma(x) = M_\Gamma^c \frac{c^x}{x!} \quad \text{for } x \in \Gamma, \tag{3.6}$$

where c is a complex balanced equilibrium of a system related to $(\mathcal{G}, \kappa)$ and M_Γ^c is a normalising constant depending on c. This is in accordance with the main result of Anderson *et al.* (2010):

Theorem 10. *Let $(\mathcal{G}, \kappa)$ be a complex balanced mass-action system. Then, there exists a unique stationary distribution on every irreducible component Γ, and it is of the form (3.6), where c is a complex balanced equilibrium of $(\mathcal{G}, \kappa)$ and M_Γ^c is a normalising constant.*

In Paper II, we study under which assumptions the converse of Theorem 10 holds, namely under what hypotheses the existence of stationary distributions of the form (3.6) for a certain c implies the existence of a complex balanced equilibrium for the related deterministic model. We prove the following:

Theorem 11. *Let $\mathcal{G}$ be an almost essential reaction network, $\kappa \in \mathbb{R}^m_+$ a vector of rate constants and $c \in \mathbb{R}^n_+$ a vector with positive entries. The probability distribution $\pi_\Gamma \colon \Gamma \to (0,1]$, defined by (3.6) is a stationary distribution for the stochastic mass-action system $(\mathcal{G}, \kappa)$ for all irreducible components $\Gamma \subseteq \mathbb{N}^n$ of $\mathcal{G}$ if and only if c is a complex balanced equilibrium for $(\mathcal{G}, \kappa)$.*

Some of the assumptions can be relaxed, and in the paper we discuss it. We also investigate whether some of the properties of complex balanced stationary distributions hold for the larger family of stationary distributions of the form 3.6. In general the answer is negative: in particular, the existence of a stationary distribution of the form 3.6 on a positive irreducible component does not imply the same form for stationary distributions on other irreducible components, as it is shown by examples in the paper.

3.3 Overview of Paper III

Even though the dynamics of the stochastic and the deterministic reaction systems are linked by the Classical Scaling and by the more general results of Ball *et al.* (2006), Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013) introduced in Section 3.1, a connection only exists for the evolution of the two models up to a finite time T. Very few results are available on the link between the behaviour of the two models when time goes to infinity. One of such results is Theorem 10: for complex balanced systems, the limit distribution of the stochastic model is a rescaled product-form Poisson distribution centred in the deterministic limit point of the deterministic model. Therefore, a concordance between the limit behaviours of the deterministic and stochastic reaction systems exists in the complex balanced case. However, such a close relation does not always exist, as discussed in Chapter 1 with reference to the reaction system

$$S_1 + S_2 \xrightarrow{\kappa_1} 2S_2 \qquad S_2 \xrightarrow{\kappa_2} S_1 \tag{3.7}$$

where for consistency with the introduced notation, we substituted A with S_1 and B with S_2. When stochastically modelled, the mass-action system hits with probability one the state $(\|X(0)\|_1, 0)$, while the solution of the deterministic mass-action system tends to $(\kappa_2/\kappa_1, \|z(0)\|_1 - \kappa_2/\kappa_1)$, provided that $z(0)$ is positive and $\|z(0)\|_1 > \kappa_2/\kappa_1$.

If the initial conditions are very large, then the behaviour of the two models is very different when the time goes to infinity. In this regards, in Anderson *et al.* (2014b) it is proven a result stating that this is a common feature for a class of reaction systems with *absolute concentration robust species* (referred to as ACR species in the rest of this manuscript). We first give the necessary definitions:

Definition 5. *Let $(\mathcal{G}, K)$ be a deterministic reaction system. We say that a species S_i possesses* absolute concentration robustness *(ACR) if for any two positive equilibria q, q' of the system, we have $q_i = q'_i$. In this case, the species S_i is called an ACR* species *and, if a positive equilibrium q exists, q_i is called an* ACR value*. If $(\mathcal{G}, K)$ possess a non empty set of ACR species, we call it an ACR system.*

Consider a deterministic reaction system that has no steady states, or only one. According to Definition 5, all the species are ACR species, however in these cases the ACR property is not particularly meaningful. ACR species are important in biochemistry because at equilibrium they always provide the same kind of response, independently of the changing of the environment and of eventual biological switches of the system. They have been observed in different biochemical cellular mechanisms including signal transduction cascades and gene regulatory networks, further details can be found for example in Shinar & Feinberg (2010), Blanchini & Franco (2011) and Karp *et al.* (2012).

Structural properties implying absolute concentration robustness have been investigated by Shinar & Feinberg (2010). To state the result, we first need the following definition:

Definition 6. *We say that two complexes $y, y' \in \mathcal{C}$* differ only in species *S if their difference is a non-zero multiple of the species S*

As an example, the complexes $S_1 + S_2$ and S_1 in (3.7) differ only in species S_2. We can now state the main result of Shinar & Feinberg (2010):

Theorem 12. *Let $\mathcal{G}$ be a reaction network which has a deficiency of one, and suppose that in the network there are two non-terminal complexes that differ only in species S. Then, S is an ACR species for any mass-action system $(\mathcal{G}, \kappa)$.*

It is not guaranteed that a positive steady state exists, and in order to state it we need to add it to the assumptions. In Anderson *et al.* (2014b) the following result is stated:

Theorem 13. *Let $\mathcal{G}$ be a reaction network which is conservative, has a deficiency of one, and for which the deterministically modelled mass-action system $(\mathcal{G}, \kappa)$ admits a positive equilibrium for some choice of κ. Suppose that there are two non-terminal*

complexes y and y' for which $y < y'$. Then, the stationary distribution of any stochastic reaction system of the form $(\mathcal{G}, K)$ is concentrated on a set of states where all the non-terminal reactions are not active.

In the case of (3.7), both the above Theorems can be applied, and an indication of the discrepancy in the limit behaviour of the deterministic and stochastic models can be derived. In this case, the states where the non-terminal reactions are not active are the states where $S_2 \to S_1$ is not active, therefore their second entry is equal to zero.

What are in general the implications of Theorems 12 and 13 in terms of limit behaviour of the stochastic and deterministic systems? In order to understand it, we now restrict ourselves to systems satisfying the hypotheses of Theorem 13, with mass-action kinetics and with the complexes y and y' differing only in species S. By Theorem 12, such systems are ACR systems, when deterministically modelled. In particular, the value of S at any positive equilibrium is constant. On the other hand, by Theorem 13, the processes of the stochastically modelled systems are absorbed by a set of states near the boundaries of $\mathbb{R}^{|\mathcal{X}|}$ (the states where the non-terminal reactions are not active) with probability one. Depending on the considered network, the counts of S may be large in the absorbing set because their size may not affect the fact that the terminal reactions are not active, as it happens for example in (3.7). This alone does not imply a discrepancy between the deterministic and the stochastic systems as observed in (3.7), but if we further assume that the positive steady states of the deterministically modelled systems are asymptotically stable, on one hand we will have an ODE soultion tending to a positive state, and on the other hand we have a stochastic process being absorbed by the boundaries, exactly as in (3.7). Stability of the steady states is not implied by the assumptions of Theorem 13, as shown by

$$S_1 + S_2 \xrightarrow{\kappa_1} 2S_2 \qquad 2S_1 + S_2 \xrightarrow{\kappa_2} 3S_1$$

whose positive steady states are unstable. In this case the deterministically modelled system is attracted by the boundaries, as the stochastically modelled one. However, stability of the steady states is expected for a large class of the systems of interest.

Hence, it is a fact that sometimes there is a discrepancy between the limit behaviour of the deterministically and stochastically modelled ACR systems. However, we do not know how long we have to wait for this discrepancy to arise. This is the setting where Paper III was conceived: in the paper the disparity between the two deterministic and stochastic models is composed by considering their behaviour up to a fixed finite time T. We consider an ACR system with at least two positive steady state, and assume that the initial condition of the stochastic reaction system is near to one steady state: in particular the values of the ACR species will be near

to their ACR values. We then let the amount of the non-ACR species increase with a parameter N, such that the initial condition is still near to a positive steady state, and study the evolution of the ACR species up to finite time T. The setting is that of the multiscale reaction systems of Ball *et al.* (2006), Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013), but in this case we are mostly interested in the behaviour of the so-called fast species, for N going to infinity. Under some additional technical assumptions, we are able to prove that, up to time T, the ACR species stay on average near to their ACR value, as it happens for the deterministic model. As an example, consider again 3.7 and consider a sequence of stochastic processes associated with the system such that $X^N(0) = (\kappa_2/\kappa_1, N - \kappa_2/\kappa_1)$. The main result in Paper III implies that in this case, for any continuous function $\widehat{g}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ with at most polynomial growth rate we have

$$\int_0^{\cdot\wedge T} \Big(\widehat{g}(X_1^N(s)) - E[\widehat{g}(J)]\Big) ds \underset{N\to\infty}{\Longrightarrow} 0,$$

where J is a Poisson random variable with parameter κ_2/κ_1 and $\Rightarrow$ denotes the weak convergence. In particular, we have

$$\int_0^{\cdot\wedge T} \left(X_1^N(s) - \frac{\kappa_2}{\kappa_1}\right) ds \underset{N\to\infty}{\Longrightarrow} 0.$$

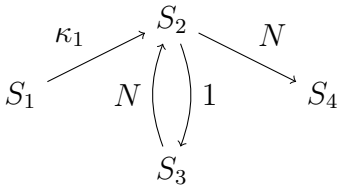
4 Perspective

In this chapter we discuss the perspectives suggested by the work done in Papers I-III, and possible future works.

4.1 Future work on intermediate species

Some questions are left open by Paper I, some of which are proposed in the Discussion Section of the paper itself. The first question concerns the generalization of the concept of intermediate species: they are defined as single species that in the reaction graph are located in a directed path connecting two non-intermediate complexes. One question is whether we could define intermediate complexes as more general linear combination of species (for example $2S_1$ or $S_1 + S_2$) still sitting in a path connecting two non-intermediate complexes, and which kind of reduction can be performed in this case. Some reductions in this more general case are shown in Examples 5.1 and 5.2 in Paper I.

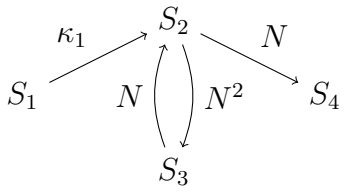
The convergence results of Paper I rely on the assumptions that some key sets of intermediate species are fast consumed. For example, consider the reaction system



with $\alpha = 0$. Here, if N is large we can expect that for most of the time no molecules of intermediate species are present, since whenever a molecule of S_2 is produced, it is most likely fast converted in a molecule of S_4. Under this assumption, we can approximate the original model by

$$S_1 \xrightarrow{\kappa_1} S_4$$

in the sense specified by Theorem 4.6 in Paper I. Now consider the same reaction network with the following different choice of rate constants:



(4.1)

Here, the time we have to wait between the production of a molecule of S_2 and the eventual production of S_4 does not tend to zero for N going to infinity: on average, the molecule of S_2 will undergo the sequence of transformations $S_2 \to S_3 \to S_2$ N times before producing a molecule of S_4, and each of this cycles of transformations will take an average time of $1/N$ to be completed. As a consequence, the results of Paper I cannot be applied (nor can be applied the results from Kang & Kurtz (2013) and Pfaffelhuber & Popovic (2013) if we want to keep track of the evolution of the species S_4). However, the waiting time from the production of a molecule of S_4 to the eventual production of a molecule of S_4 is distributed as

$$\tau^N = \sum_{j=1}^{G^N} (E_j^{1,N} + E_j^{2,N}),$$

where G^N is a geometric random variable with parameter $1/N$ expressing the number of cycles $S_2 \to S_3 \to S_2$ we observe before the production of S_4, $E_j^{1,N}$ is an exponential random variable with parameter $N^2 + N$ which is the waiting time for the reaction $S_2 \to S_3$ to occur (given that it will happen before $S_2 \to S_4$), and finally $E_j^{2,N}$ is an exponential random variable with parameter N which is the waiting time for the reaction $S_3 \to S_2$ to occur. All the random variables G^N, $(E_j^{1,N})_j$ and $(E_j^{2,N})_j$ are independent. At this point, it is not hard to argue that the increment given to τ^N by the random variables $(E_i^{1,N})_i$ tends to zero almost surely for N going to infinity. Moreover, by considering the characteristic functions we have

$$\begin{aligned} E\left[e^{it\sum_{j=1}^{G^N} E_j^{2,N}}\right] &= E\left[\prod_{j=1}^{G^N} e^{itE_j^{2,N}}\right] = \sum_{q=1}^{\infty}\prod_{j=1}^{q} E\left[e^{itE_j^{2,N}}\right] P\left(G^N = q\right) \\ &= \frac{1}{N-1}\sum_{q=1}^{\infty}\left(\frac{N-1}{N-it}\right)^q = \frac{1}{1-it}. \end{aligned}$$

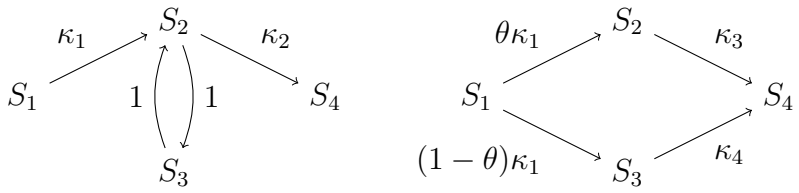
Since the last term is the characteristic function of an exponential random variable with parameter 1, we have that τ^N converges in distribution to it. Therefore, the

dynamics of S_1 and S_2 can be approximated by means of the simplified reaction system

$$S_1 \xrightarrow{\kappa_1} D \xrightarrow{1} S_4$$

where D is a fictitious species. A proper computational and analytical analysis of the simplified system can be easily performed since the large constant rates have been eliminated. Studying this kind of approximation technique under general assumptions could be an interesting line of research for future works.

Another interesting research direction concerns identifiability of the intermediate structure. This is better understood with an example: consider the two mass-action systems



(4.2)

where κ_1, κ_2, κ_3 and κ_4 are positive real numbers and θ is a real number in $(0, 1)$. Suppose that the initial conditions coincide for the two systems, and that the initial amount of the intermediate species S_2 and S_3 is zero. Then, for any positive κ_1 and κ_2 there exists a choice of parameters κ_3, κ_4 and θ such that the evolutions of the species S_1 and S_4 in the two stochastic mass-action systems have exactly the same distribution. Moreover, for the same choice of parameters the trajectories of the species S_1 and S_4 coincide in the two systems when deterministically modelled, as well. The choice of parameters is given by

$$\begin{aligned}
\kappa_3 &= \frac{2 + \kappa_2 + \sqrt{4 + \kappa_2^2}}{2} \\
\kappa_4 &= \frac{2 + \kappa_2 - \sqrt{4 + \kappa_2^2}}{2} \\
\theta &= \frac{\kappa_2^2 + \kappa_2\sqrt{4 + \kappa_2^2}}{4 + \kappa_2^2 + (2 + \kappa_2)\sqrt{4 + \kappa_2^2}}\,.
\end{aligned}$$

In the deterministic case, this choice generates in the two systems the same *delay*, as defined in Appendix IA (see equation (IA.6)). Moreover, the consumption rate of S_1 in the two models does not change.

The example shows that if we can only observe the evolution of non-intermediate species, then we may not be able to distinguish between two different models, both in the stochastic and in the deterministic settings. The questions we would like

to address concern general conditions under which the unidentifiability of the intermediate structure holds, and more specifically whether, given the dynamics of non-intermediate species, we can always choose a canonical minimal intermediate structure without cycles. This work could be theoretical relevant, since often intermediate species are impossible to detect in controlled experiments. Moreover, it could provide more means to simplify reaction systems with intermediate species. Indeed, we could be able to transform a system with fast consumed intermediate species into a system with no high rates, therefore easier to analyse. Moreover, it could provide us with a tool to easily find approximated models as the one discussed for (4.1). Finally, we might be able to transform intermediate structures with cycles into more tractable intermediate structures without cycles, as it happens for (4.2).

Another interesting perspective suggested by Paper I concerns a possible link between the limit behaviour of the original model with intermediate species and that of the reduced model. This is better explained in the next section.

4.2 Future work on stationary distribution

Except for complex balanced reaction systems (Jahnke & Huisinga, 2007; Anderson *et al.*, 2010) and for system whose process can be identified as a birth-death process, there are no means to explicitly calculate the stationary distribution of a reaction system. In fact, it may often be a hard task to simply postulate the existence of a stationary distribution, when the cardinality of the states that can be visited by the process is infinite. Any new result in this direction can be helpful, we list here some open question:

i) In the context of Paper I, does the existence of a stationary distribution for the original model imply the existence of a stationary distribution for the reduced system? Moreover, we have shown in Example 5.4 in Paper I that the stationary distribution of the original model may differ greatly from that of the reduced system. Under which additional assumptions do we have convergence of the stationary distributions?

ii) Is the conjecture stated at the end of Paper II true? Namely, if a mass-action system has a rescaled product-form Poisson stationary distribution on one irreducible component Γ for any choice of rate constants, is there a complex balanced mass-action system whose evolution on Γ has the same distribution as for the original process?

iii) Is it possible to generalize the results in Paper II and find broader sufficient conditions implying product-form stationary distributions? Are there necessary conditions?

iv) Is it possible to generalize Theorem 21 in Paper II and prove general results on the form of the irreducible components of a network?

v) Under which assumptions are the stationary distribution of a stochastic reaction system and the limit behaviour of a corresponding deterministic system related?

Answering the questin of point i may help in finding the stationary distribution of a reaction system by studying that of a reduced one. Note that a similar question may be interesting also in the deterministic case (see Section "Limit Behaviour" in Appendix IA): under which additional assumptions the limit behaviour of the reduced deterministic reaction system approximates that of the original deterministic reaction system?
The question of point (ii) is in the spirit of Thereom 2: we want to know whether the fact that a certain property holds for any choice of rate constants implies something on the reaction network.
In general, answering any of the above questions may provide important information on the stationary distribution of a stochastic reaction system, at least for a class of them.

Another perspective, suggested by Paper III, concerns the case when the stationary distribution of a stochastic mass-action system and the limit behaviour of the corresponding deterministic system differ greatly. In such a case, what is the typical time at which we can observe the discrepancy? In Paper III we proved that, if we let the amount of the non-ACR molecules tend to infinity in the initial condition, then up to a fixed finite time T there is a substantial concordance between the behaviours of the stochastic and the deterministic models. Does the same hold if we let T go to infinity as a function of the initial conditions? If so, does it hold for a more general class of reaction systems than that introduced in Paper III?

A study on quasi-stationary distributions, or more in general quasi-stationary phases of the stochastic systems would be also of interest. In particular, it could be interesting to understand when a quasi-stationary regime exists in a stochastic reaction system, when it appears and for how long it can be observed. A related question that still has not received much interest in the CRNT literature concerns the *mixing times* of stochastic reaction networks, and whether their expected value is linked to structural properties of the reaction network.

4.3 Future work on ACR species

In Shinar & Feinberg (2010), sufficient conditions for a species to be ACR have been found. However, these conditions are far from including all the ACR systems, so

further work in this direction might be of interest, as well as providing necessary conditions for a species to be ACR. Moreover, the result in Anderson *et al.* (2014b) is stated for ACR systems that fall in the assumptions of Shinar & Feinberg (2010), hence we might wonder whether it holds for a broader class of ACR systems.

Bibliography

Anderson, DF, & Kurtz, TG. 2011. Continuous time Markov chain models for chemical reaction networks. *Pages 3–42 of:* Koeppl, Heinz, Densmore, Douglas, Setti, Gianluca, & Di Bernardo, Mario (eds), *Design and analysis of biomolecular circuits: Engineering approaches to systems and synthetic biology*. Springer.

Anderson, DF, & Kurtz, TG. 2015. *Stochastic analysis of biochemical systems*. Springer.

Anderson, DF, Craciun, G, & Kurtz, TG. 2010. Product-form stationary distributions for deficiency zero chemical reaction networks. *Bulletin of mathematical biology*, **72**(8), 1947–1970.

Anderson, DF, Craciun, G, Gopalkrishnan, M, & Wiuf, C. 2014a. *Lyapunov functions, stationary distributions, and non-equilibrium potential for chemical reaction networks*. arXiv:1410.4820.

Anderson, DF, Enciso, G, & Johnston, MD. 2014b. Stochastic analysis of biochemical reaction networks with absolute concentration robustness. *Journal of the royal society interface*, **11**(93), 20130943.

Anderson, RM, & May, RM. 1992. *Infectious diseases of humans: dynamics and control*. Vol. 28. Wiley Online Library.

Angeli, D, De Leenheer, P, & Sontag, ED. 2007. A Petri net approach to the study of persistence in chemical reaction networks. *Mathematical biosciences*, **210**(2), 598–618.

Ball, K, Kurtz, TG, Popovic, L, & Rempala, G. 2006. Asymptotic analysis of multiscale approximations to reaction networks. *The annals of applied probability*, **16**(4), 1925–1961.

Blanchini, F, & Franco, E. 2011. Structurally robust biological networks. *Bmc systems biology*, **5**(1), 74.

Boros, B. 2013. On the existence of the positive steady states of weakly reversible deficiency-one mass action systems. *Mathematical biosciences*, **245**(2), 157–170.

Cornish-Bowden, A. 2004. *Fundamentals of enzyme kinetics*. Third edn. London: Portland Press.

Craciun, G. 2015. Toric differential inclusions and a proof of the global attractor conjecture. *arxiv preprint arxiv:1501.02860*.

Craciun, G, Dickenstein, A, Shiu, A, & Sturmfels, B. 2009. Toric dynamical systems. *Journal of symbolic computation*, **44**(11), 1551–1565.

Cummings, R, Doty, D, & Soloveichik, D. 2014. Probability 1 computation with chemical reaction networks. *Pages 37–52 of: DNA computing and molecular programming*. Springer.

Doty, D, Lutz, JH, Patitz, MJ, Schweller, RT, Summers, SM, & Woods, D. 2012. The tile assembly model is intrinsically universal. *Pages 302–310 of: Foundations of computer science (FOCS), 2012 IEEE 53rd annual symposium on*. IEEE.

Érdi, P, & Tóth, J. 1989. *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*. Manchester University Press.

Ethier, SN, & Kurtz, TG. 2009. *Markov processes: characterization and convergence*. Vol. 282. John Wiley & Sons.

Feinberg, M. 1972. Complex balancing in general kinetic systems. *Archive for rational mechanics and analysis*, **49**(3), 187–194.

Feinberg, M. 1979. Lectures on chemical reaction networks. *Notes of lectures given at the mathematics research center, university of wisconsin*.

Feinberg, M. 1987. Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems. *Chemical engineering science*, **42**(10), 2229–2268.

Feinberg, M. 1988. Chemical reaction network structure and the stability of complex isothermal reactors—ii. multiple steady states for networks of deficiency one. *Chemical engineering science*, **43**(1), 1–25.

Feliu, E. 2015. Injectivity, multiple zeros and multistationarity in reaction networks. *Page 20140530 of: Proceedings of the royal society of london a: Mathematical, physical and engineering sciences*, vol. 471. The Royal Society.

Feliu, E, & Wiuf, C. 2013. Simplifying biochemical models with intermediate species. *Journal of the royal society interface*, **10**(87).

Gunawardena, J. 2012. Some lessons about models from Michaelis and Menten. *Molecular biology of the cell*, **23**(4), 517.

Gupta, A, & Khammash, MH. 2014. Determining the long-term behavior of cell populations: A new procedure for detecting ergodicity in large stochastic reaction networks. *Pages 1711–1716 of: World congress*, vol. 19.

Gusak, D, Kukush, A, Kulik, A, Mishura, Y, & Pilipenko, A. 2010. *Theory of stochastic processes: With applications to financial mathematics and risk theory.* Springer New York.

Horn, F. 1972. Necessary and sufficient conditions for complex balancing in chemical kinetics. *Archive for rational mechanics and analysis*, **49**(3), 172–186.

Horn, F. 1974. The dynamics of open reaction systems. *Pages 125–137 of: Mathematical aspects of chemical and biochemical problems and quantum chemistry (proc. siam-ams sympos. appl. math., new york, 1974)*, vol. 8.

Horn, F, & Jackson, R. 1972. General mass action kinetics. *Archive for rational mechanics and analysis*, **47**(2), 81–116.

Ingram, PJ, Stumpf, MPH, & Stark, J. 2008. Nonidentifiability of the source of intrinsic noise in gene expression from single-burst data. *PLoS computational biology*, **4**(10).

Jahnke, T, & Huisinga, W. 2007. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of mathematical biology*, **54**(1), 1–26.

Joshi, B. 2015. A detailed balanced reaction network is sufficient but not necessary for its Markov chain to be detailed balanced. *Discrete and continuous dynamical systems - series b*, **20**(4), 1077–1105.

Joshi, B, & Shiu, A. 2013. Atoms of multistationarity in chemical reaction networks. *Journal of mathematical chemistry*, **51**(1), 153–178.

Kang, H-W, & Kurtz, TG. 2013. Separation of time-scales and model reduction for stochastic reaction networks. *The annals of applied probability*, **23**(2), 529–583.

Karlebach, G, & Shamir, R. 2008. Modelling and analysis of gene regulatory networks. *Nature reviews molecular cell biology*, **9**(10), 770–780.

Karp, RL, Millán, MP, Dasgupta, T, Dickenstein, A, & Gunawardena, J. 2012. Complex-linear invariants of biochemical networks. *Journal of theoretical biology*, **311**, 130–138.

Kermack, WO, & McKendrick, AG. 1927. A contribution to the mathematical theory of epidemics. *Pages 700–721 of: Proceedings of the royal society of london a: Mathematical, physical and engineering sciences*, vol. 115. The Royal Society.

Kurtz, TG. 1972. The relationship between stochastic and deterministic models for chemical reactions. *The journal of chemical physics*, **57**(7), 2976–2978.

Kurtz, TG. 1978. Strong approximation theorems for density dependent markov chains. *Stochastic processes and their applications*, **6**(3), 223–240.

Kurtz, TG. 1992. Averaging for martingale problems and stochastic approximation. *Pages 186–209 of: Applied stochastic analysis*. Springer.

Mairesse, J, & Nguyen, H-T. 2009. Deficiency zero Petri nets and product form. *Pages 103–122 of: Applications and theory of Petri nets*. Springer.

Markevich, NI, Hoek, JB, & Kholodenko, BN. 2004. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *The journal of cell biology*, **164**(3), 353–359.

May, RM. 2001. *Stability and complexity in model ecosystems*. Vol. 6. Princeton University Press.

Müller, S, Feliu, E, Regensburger, G, Conradi, C, Shiu, A, & Dickenstein, A. 2013. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Foundations of computational mathematics*, 1–29.

Murray, JD. 2002. *Mathematical biology I: An introduction, vol. 17 of interdisciplinary applied mathematics*.

Norris, JR. 1998. *Markov chains*. Cambridge university press.

Paulevé, L, Craciun, G, & Koeppl, H. 2014. Dynamical properties of discrete reaction networks. *Journal of mathematical biology*, **69**(1), 55–72.

Peschel, M, & Breitenecker, F. 1984. Socio-economic consequences of the Volterra approach. *Cybernetics and systems research*, **2**.

Peschel, M, & Mende, W. 1986. *The predator-prey model: do we live in a Volterra world?* Springer.

Pfaffelhuber, P, & Popovic, L. 2013. Scaling limits of spatial chemical reaction networks. *arxiv preprint arxiv:1302.0774*.

Rothemund, PWK, & Winfree, E. 2000. The program-size complexity of self-assembled squares. *Pages 459–468 of: Proceedings of the thirty-second annual ACM symposium on Theory of computing*. ACM.

Shinar, G, & Feinberg, M. 2010. Structural sources of robustness in biochemical reaction networks. *Science*, **327**(5971), 1389–1391.

Soloveichik, D, Seelig, G, & Winfree, E. 2010. DNA as a universal substrate for chemical kinetics. *Proceedings of the national academy of sciences*, **107**(12), 5393–5398.

Thomson, M, & Gunawardena, J. 2009. Unlimited multistability in multisite phosphorylation systems. *Nature*, **460**(7252), 274–277.

Weidlich, W, & Haag, G. 2012. *Concepts and models of a quantitative sociology: the dynamics of interacting populations*. Vol. 14. Springer Science & Business Media.

Papers

I

Elimination of Intermediate Species in Multiscale Stochastic Reaction Networks

Daniele Cappelletti
Department of Mathematical Sciences
University of Copenhagen

Carsten Wiuf
Department of Mathematical Sciences
University of Copenhagen

Publication details

Submitted (2014).

Elimination of Intermediate Species in Multiscale Stochastic Reaction Networks

Daniele Cappelletti* Carsten Wiuf*

Abstract

We study networks of biochemical reactions modelled by continuous-time Markov processes. Such networks typically contain many molecular species and reactions and are hard to study analytically as well as by simulation. Particularly, we are interested in reaction networks with intermediate species such as the substrate-enzyme complex in the Michaelis-Menten mechanism. Such species are virtually in all real-world networks, they are typically short-lived, degraded at a fast rate and hard to observe experimentally.

We provide conditions under which the Markov process of a multiscale reaction network with intermediate species is approximated by the Markov process of a simpler reduced reaction network without intermediate species. We do so by embedding the Markov processes into a one-parameter family of processes, where reaction rates and species abundances are scaled in the parameter. Further, we show that there are close links between these stochastic models and deterministic ODE models of the same networks.

1 Introduction

Reliable mathematical models of biochemical reaction networks are of great interest for the analysis of experimental data and theoretical biochemistry. Such models can provide qualitative information on biochemical systems as well as provide means to simulate networks and to estimate unknown parameters. The classical stochastic model of a reaction network is a continuous-time Markov process, where the states are configurations of species numbers and the transitions are changes caused by reactions. We refer to this Markov process as a *stochastic reaction network (SRN)*. Unfortunately the set of reactions and chemical species is often very large, and the related Markov process is too complicated to be studied analytically or by modern computers. Thus, the necessity of simplifying the full model arises. Perhaps the first result in this direction is due to Kurtz [1972], where a deterministic weak limit for stochastic reaction networks is obtained [see also Kurtz, 1977/78]. More recently, in Ball et al. [2006], Kang and Kurtz [2013], Pfaffelhuber and Popovic [2013], similar asymptotic results have been obtained under more general scaling conditions than those applied in Kurtz [1972, 1977/78]. Here the limit might have stochastic as well as deterministic components, and the limit network might consist of simplified reactions with fewer species. In this context the concept of *model reduction* arises naturally.

A famous and well studied example of a biochemical system is the Michaelis-Menten mechanism for enzyme kinetics [Cornish-Bowden, 2004, Kang and Kurtz, 2013, Härdin et al., 2009, Thomas et al., 2012, Rao and Arkin, 2003]. It is described by the reactions

$$E + R \rightleftarrows H \longrightarrow E + P$$

where E denotes an enzyme, R a reacting substrate and P a product. H is an intermediate, or transient, species formed by E and R, and it is usually unstable. Whenever a reaction occurs, say $E + R \to H$, then the number of molecules changes accordingly, that is, the numbers of E and R molecules are each reduced by one, while the number of H molecules is increased by one.

If we assume that at least one of the reactions $H \to E + R$ and $H \to E + P$ is so fast that a produced molecule of H is quickly degraded before any other reaction takes place (that is, at any time at most

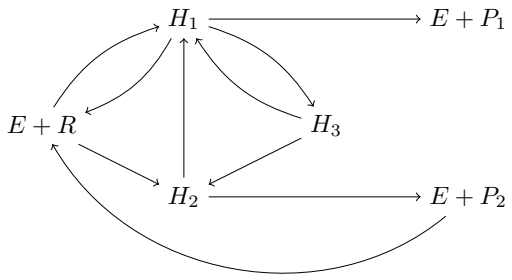
*Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark. The authors are supported by funding from the Carlsberg Foundation, the Lundbeck Foundation, the Danish Cancer Society the Danish Research Councils.

one molecule of H is present), then it seems reasonable that the Markov process could be approximated by a simpler Markov process, corresponding to the *reduced* reaction network

$$E + R \longrightarrow E + P$$

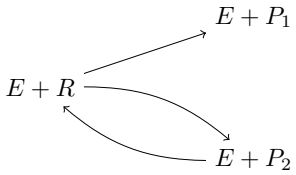
where the reaction rate is determined from the original reaction rates. Intuitively, the rate is the rate of $E + R \to H$ multiplied by the probability that the reaction $H \to E + P$ occurs instead of $H \to E + R$. Under this reduction the number of enzyme molecules E becomes constant. In essence, we are here dealing with *time-scale separation*, in addition to *species elimination* and *dimensionality reduction* (both in terms of the number of reactions as well as the number of species).

Another, perhaps more interesting example, is the following reaction network:



(1.1)

It describes the catalytic transformation of a species R into the species P_1 or P_2, through a chain of intermediate steps, denoted by the species H_1, H_2 and H_3. Whenever the reaction $E+R \to H_1$ occurs, a sequence of reactions between intermediate species will take place (for example, $H_1 \to H_3 \to H_1$) before a final complex is produced, such as $E + P_1$. If the time spent in intermediate states is small, we might approximate the reaction paths proceeding through the formation and quick degradation of intermediate species by direct reactions. In other words, it is reasonable to contract reaction paths passing through any intermediate species to obtain



for a suitable choice of reaction rates. Note that there is an infinite number of such reaction paths. We will provide conditions that guarantee that the original SRN can be well approximated, in a certain sense, by the reduced SRN, or more accurately, that the Markov process describing the original system is well approximated by the Markov process of the reduced system.

For this aim, we introduce a family of *kinetics* (reaction rates) indexed by a parameter N and study the relationship between the original and the reduced SRNs as $N \to \infty$. The analysis builds on the previous work Feliu and Wiuf [2013a] [see also Feliu and Wiuf, 2012, 2013b], as well as on Ball et al. [2006], Kang and Kurtz [2013], Pfaffelhuber and Popovic [2013]. In Feliu and Wiuf [2013a], a mathematical framework is developed for the elimination of intermediate species in deterministically modelled reaction networks, using ODEs. Properties of the steady states in the original ODE system are related to similar properties of the steady states in the reduced ODE system by means of a formal relationship between the original and the reduced network. Here we are not concerned about the steady states nor about the equilibrium distributions of SRNs, but about the trajectories of SRNs up to a finite fixed time $T > 0$. Our aim is to approximate the dynamics of the original system with intermediate species by means of the dynamics of a simplified model, where intermediate species are eliminated. Though we arrive at our reduced model through a different route than Feliu and Wiuf [2013a], we will show that there are close links to ODE models and that our reduced network in fact is that of Feliu and Wiuf [2013a].

We will study different types of convergence of stochastic processes associated with SRNs as $N \to \infty$. The limit is taken assuming that the consumption rates (at least some of them) of the intermediate

species approach infinity according to N. Also the molecular abundances might be scaled in powers of N in the spirit of the *multiscale analysis* performed in Ball et al. [2006], Kang and Kurtz [2013], Pfaffelhuber and Popovic [2013]. These papers deal with various forms of model reduction. However, the elimination of intermediate species we aim to achieve is not possible in these settings. On the other hand, our approximating model might in some cases be further reduced by techniques developed in these papers, hence our approach might be considered complementary to theirs.

2 Preliminaries and definitions

The space of real (natural) vectors with entries indexed by a finite set A is denoted by $\mathbb{R}^A$ ($\mathbb{N}^A$), and for any vector $v \in \mathbb{R}^A$ ($\mathbb{N}^A$), we denote the entry corresponding to $a \in A$ by $v(a)$. Moreover, for any two vectors $v, w \in \mathbb{R}^A$ ($\mathbb{N}^A$) we write $v > w$ if the inequality holds component-wise. Furthermore, $|v|$ denotes the usual Euclidean norm of v. Finally, if A is a finite set, we let $\#A$ denote the cardinality of A. Given two real numbers x, y, we will often use the notation $x \vee y$ or $x \wedge y$ to denote the maximum and the minimum of x and y, respectively.

A reaction network consists of a set of species $\mathcal{X}$, a set of complexes $\mathcal{C}$, and a set of reactions $\mathcal{R}$. Formally, $\mathcal{X}$ is a finite non-empty set $\{S_1, S_2, \dots, S_n\}$, $\mathcal{C} = \{y_1, y_2, \dots, y_m\}$ is a non-empty set of non-negative linear combinations of elements of $\mathcal{X}$ and $\mathcal{R}$ is a finite non-empty subset of $\mathcal{C} \times \mathcal{C}$, such that $(y_i, y_i) \notin \mathcal{R}$ for all i. We identify $\mathcal{X}$ and $\mathcal{C}$ with finite subsets of $\mathbb{N}^{\mathcal{X}}$. If $(y_i, y_j) \in \mathcal{R}$ we write $y_i \to y_j$ and we say that y_i is the *reactant* and y_j is the *product*. Throughout the paper we will denote an object O associated with a reaction $r \colon y_i \to y_j$ by O_r or O_{ij} indifferently. Furthermore, for each reaction $r \colon y_i \to y_j \in \mathcal{R}$, we define the *reaction vector*

$$\xi_r = y_j - y_i.$$

For further background on reaction networks, see Érdi and Tóth [1989], Anderson and Kurtz [2011].

A complex $y \in \mathcal{C}$ is given as $y = (y(S_1), \dots, y(S_n))$ and $y(S)$ is called the *stoichiometric coefficient* of the species S in y. Furthermore, we define the *support* of y as the set of species S such that $y(S) > 0$, in which case we write $S \in y$. Moreover, define $\mathcal{C}_S$ as the complexes whose support contains S and $\mathcal{R}_S$ as the reactions in $\mathcal{R}$ that change the counts of S:

$$\mathcal{C}_S = \{y \in \mathcal{C} : \ S \in y\}, \tag{2.1}$$

$$\mathcal{R}_S = \{r \in \mathcal{R} : \ \xi_r(S) \neq 0\}. \tag{2.2}$$

Finally, we define a *kinetics* $\mathcal{K}$ as a set of functions indexed by $\mathcal{R}$ of the form

$$\begin{array}{rccc} \lambda_r \colon & \mathbb{N}_{\geq 0}^{\mathcal{X}} & \to & \mathbb{R}_{\geq 0} \\ & x & \mapsto & \lambda_r(x). \end{array}$$

Intuitively, λ_r is the rate by which reaction r occurs and it will be referred to as the *reaction rate*. We allow reaction rates to be constantly 0, in which case the corresponding reaction could be removed from the network.

A reaction network equipped with a kinetics can be modelled as a continuous-time Markov process $X_\cdot$ on $\mathbb{N}^{\mathcal{X}}$, where $X_t(S)$ is the number of molecules of the species S at time t. Taken together with $\mathcal{K}$ and $X_\cdot$, a reaction network is called a *stochastic reaction network* (*SRN*). The state of $X_\cdot$ changes whenever a reaction takes place, for example, if the reaction r occurs at time t^* the new state is

$$X_{t^*} = X_{t^*-} + \xi_r.$$

The kinetics $\mathcal{K}$ represents the transition rates for the process $X_\cdot$, such that

$$X_t = X_0 + \sum_{r \in \mathcal{R}} \xi_r Y_r \left(\int_0^t \lambda_r(X_s) ds \right), \tag{2.3}$$

with $Y_r(\cdot)$ independent and identically distributed unit-rate Poisson processes [Kurtz, 1977/78]. The random variable $Y_r \left(\int_0^t \lambda_r(X_s) ds \right)$ counts how many times the reaction r has occurred up to time t. This

stochastic model is typically chosen if the number of reactant molecules is low, so that the behaviour of the system is similar to the evolution of a jump process. Changes occur only in a discrete set of time points and it is uncertain which reaction will take place next.

A typical choice of kinetics is *mass-action kinetics*, where the reaction rate of $r\colon y_i \to y_j$ is given by

$$\lambda_r(x) = k_r \prod_{S\in y_i} \frac{x(S)!}{(x(S)-y_i(S))!}\mathbb{1}_{\{x(S)\geq y_i(S)\}},$$

and k_r are non-negative real numbers, called *rate constants*. We usually express this as $y_i \xrightarrow{k_r} y_j$. Note that the reaction rates are proportional to the number of ordered subsets of molecules that can give rise to an occurrence of the reaction. This choice of kinetics is natural if we assume the system is well stirred.

To define a reduced reaction network we introduce the concept of an intermediate species [Feliu and Wiuf, 2013a].

Definition 2.1. Let $(\mathcal{X},\mathcal{C},\mathcal{R})$ be a reaction network and $\mathcal{V}\subset\mathcal{X}$. We say that the species in $\mathcal{V}$ are *intermediate species* (or simply *intermediates*) if the following conditions hold:

- for each $H\in\mathcal{V}$ and $y\in\mathcal{C}$, if H is in the support of y, then $y=H$. This implies that $\mathcal{V}\subset\mathcal{C}$.
- for each $H\in\mathcal{V}$, there is a directed path of complexes such that

  $$y_i \to H_{\ell_1} \to \cdots \to H \to \cdots \to H_{\ell_k} \to y_j$$

  for some complexes $y_i, y_j \in \mathcal{C}\setminus\mathcal{V}$ and $H_{\ell_i}\in\mathcal{V}$ for all $1\leq i\leq k$. The path

  $$H_{\ell_1} \to \cdots \to H \to \cdots \to H_{\ell_k}$$

  is called a *chain of intermediates*.

According to the definition, intermediate species always appear alone and with stoichiometric coefficient one. For example, the species H in the Michaelis-Menten mechanism and the species H_1, H_2 and H_3 in (1.1) meet Definition 2.1. We denote by $\mathcal{U}$, $\mathcal{W}$ the subsets of $\mathcal{C}$ such that

- for all $y\in\mathcal{U}$, there exists $H\in\mathcal{V}$, such that $y\to H\in\mathcal{R}$
- for all $y\in\mathcal{W}$, there exists $H\in\mathcal{V}$, such that $H\to y\in\mathcal{R}$

We refer to $\mathcal{U}$ and to $\mathcal{W}$, respectively, as the *initial reactants* and the *final products*. In general the two sets can have non-empty intersection (as in Example 3.1). For any initial reactant y_i we introduce the set $\mathcal{V}_i$ of intermediate species H such that $y_i\to H\in\mathcal{R}$. We index the set $\mathcal{V}$ using the ordering of the set $\mathcal{C}$, such that $H_\ell = y_\ell$ for any intermediate $H_\ell\in\mathcal{V}$. Further, we introduce the index sets U, V, V_i and W of $\mathcal{U}$, $\mathcal{V}$, $\mathcal{V}_i$ and $\mathcal{W}$, respectively, such that

$$\mathcal{U}=\{y_i\}_{i\in U}\,,\quad \mathcal{V}=\{H_\ell\}_{\ell\in V}\,,\quad \mathcal{V}_i=\{H_\ell\}_{\ell\in V_i}\,,\quad \mathcal{W}=\{y_j\}_{j\in W}\,.$$

3 The Reduced Stochastic Reaction Network

Let $(\mathcal{X},\mathcal{C},\mathcal{R})$ be a reaction network equipped with a kinetics $\mathcal{K}$ and let $\mathcal{V}\subset\mathcal{X}$ be a set of intermediate species.

The reduced reaction network obtained from $(\mathcal{X},\mathcal{C},\mathcal{R})$ is the triple

$$(\mathcal{X}\setminus\mathcal{V}\,,\,\mathcal{C}\setminus\mathcal{V}\,,\,\mathcal{R}^*), \tag{3.1}$$

where $\mathcal{R}^*$ consists of the reactions in $\mathcal{R}$ not involving intermediates and the reactions $y_i\to y_j$, where y_j is obtainable from y_i through a chain of intermediate species of $(\mathcal{X},\mathcal{C},\mathcal{R})$, as in Definition 2.1. Thus, the intermediate species have been eliminated from the original network by contraction of reaction paths.

If $(\mathcal{X},\mathcal{C},\mathcal{R})$ is equipped with a kinetics $\mathcal{K}$, then $(\mathcal{X}\setminus\mathcal{V},\mathcal{C}\setminus\mathcal{V},\mathcal{R}^*)$ inherits a kinetics $\mathcal{K}^*$ from $(\mathcal{X},\mathcal{C},\mathcal{R})$ if certain additional conditions are fulfilled. To define $\mathcal{K}^*$ we first make the following assumption:

Assumption 1 (Rate functions and intermediates)**.** The consumption of the intermediate species is governed by mass-action kinetics, that is for any $\ell, \ell' \in V$ and $j \in W$,

$$\lambda_{\ell j}(x) = k_{\ell j} x(H_\ell), \quad \text{and} \quad \lambda_{\ell \ell'}(x) = k_{\ell \ell'} x(H_\ell),$$

for some non-negative constants $k_{\ell j}$, $k_{\ell \ell'}$. This condition implies that any molecule of an intermediate species will be consumed at a constant rate. Further, we assume that all other reaction rates do not depend on H_ℓ.

Let $X_\cdot$ be the process associated with $(\mathcal{X}, \mathcal{C}, \mathcal{R})$. We enlarge the filtration of $X_\cdot$ by the σ-algebras σ_t, such that σ_t contains the information on the evolution up to time t of every occurrence of a molecule of an intermediate species in the experiment. In particular, we introduce a Markov process, that describes the dynamics, or fate, of a molecule of an intermediate species. Consider the n-th reaction occurring in $X_\cdot$ that turns a non-intermediate complex into an intermediate species. Let this reaction be $y_i \to H_\ell$ and assume it takes place at time t_n. The intermediate molecule H_ℓ will eventually be transformed into a final product y_j. The chain of transformations leading to y_j can be described by a continuous-time Markov chain $C_n(\cdot)$, starting at time t_n, with state space $\mathcal{V} \cup \mathcal{W}$ and $C_n(t_n) \in \mathcal{V}$. The final products are treated as absorbing states for the Markov process. The transition rate matrix, which is independent on n, has the following block structure:

$$Q = \left[\begin{array}{c|c} Q_{V,V} & Q_{V,W} \\ \hline 0 & 0 \end{array}\right], \tag{3.2}$$

where

$$\begin{aligned} q_{\ell \ell'} &= k_{\ell \ell'} && \text{for all } \ell, \ell' \in V \text{ with } \ell \neq \ell' \\ q_{\ell j} &= k_{\ell j} && \text{for all } \ell \in V \text{ and } j \in W \\ q_{\ell \ell} &= -\sum_{\ell' \in V} k_{\ell \ell'} - \sum_{j \in W} k_{\ell j} && \text{for all } \ell \in V. \end{aligned}$$

We define by τ_n the time until the production of the final product, i.e.

$$\tau_n = \inf\{t \geq t_n \colon C_n(t) \in \mathcal{W}\} - t_n,$$

and for all $\ell \in V_i$, we define by $\pi_{\ell j}$ the probability that the final product produced is y_j, given that the intermediate chain started in H_ℓ. Namely,

$$\pi_{\ell j} = P\left(C_n(t_n + \tau_n) = y_j \,|C_n(t_n) = H_\ell\right), \tag{3.3}$$

with $\pi_{\ell j} = 0$ if $j \notin W$. Since $C_n(\cdot)$ is a finite state Markov process with absorbing states, τ_n is almost surely finite. Moreover note that $\pi_{\ell j}$ does not depend on n, since Q does not depend on n. In this context, we have

$$\sigma_t = \sigma\left(X_s, C_n(s) \colon s \in [0,t], n \in \mathbb{N}\right). \tag{3.4}$$

Let $\mathcal{K}$ be a kinetics fulfilling Assumption 1. If we let $\lambda_{i\ell} = 0$ whenever $y_i \to H_\ell \notin \mathcal{R}$, then the kinetics $\mathcal{K}^*$ of the reduced reaction network is defined by

$$\lambda^*_{ij}(x) = \lambda_{ij}(x) + \sum_{\ell \in V_i} \pi_{\ell j} \lambda_{i\ell}(x), \tag{3.5}$$

for any $y_i \to y_j \in \mathcal{R}^*$. Thus, the rate of a reaction originating from a chain of intermediates is the sum of the rates $\lambda_{i\ell}(\cdot)$ by which the first intermediate is produced from y_i multiplied by the probability $\pi_{\ell j}$ that the chain ends in y_j. To this we add $\lambda_{ij}(\cdot)$ if the reaction $y_i \to y_j$ is already in $\mathcal{R}$.

Our main goal is to prove that the behaviour of $X_\cdot$, under certain conditions, is captured by the behaviour of the process associated with the reduced SRN. In the broader setting of multiscale models [Ball et al., 2006, Kang and Kurtz, 2013, Pfaffelhuber and Popovic, 2013], we prove that a suitable rescaled version of $X_\cdot$ can be approximated by a similarly rescaled version of the process of the reduced

SRN. We will show this by constructing a particular process $Z_\cdot$ on the same probability space as $X_\cdot$, which is distributed as the process associated with the reduced SRN, and by further proving convergence in probability of the difference between the rescaled versions of $X_\cdot$ and $Z_\cdot$ in various senses. Specifically, we are able to prove uniform punctual convergence in probability to zero as well as convergence in occupation measure (cf. Theorems 4.3 and 4.7). Under additional assumptions, we prove convergence in probability to zero of the difference of the rescaled processes in the Skorohod topology (cf. Theorems 4.5 and 4.7).

The reduced reaction network defined here is the same as the reduced reaction network introduced in Feliu and Wiuf [2013a]. Moreover, the procedure to obtain the kinetics of the reduced model coincides with that in Feliu and Wiuf [2013a]. We prove this in Theorem 3.1. It is worth noting, however, that the aims of Feliu and Wiuf [2013a] and this paper are very different. Indeed, we study various convergences of the stochastic processes associated with $(\mathcal{X}, \mathcal{C}, \mathcal{R})$, while in Feliu and Wiuf [2013a] the reaction networks are deterministically modelled through a system of ODEs, and a relation between the steady states of the original and the reduced models is investigated.

In Feliu and Wiuf [2013a], the kinetics of the reduced reaction network is given by

$$\widetilde{\lambda}_{ij}(x) = \lambda_{ij}(x) + \sum_{\ell \in V_i} k_{\ell j} \mu_{i\ell}(x), \tag{3.6}$$

where $\mu_{i\ell}$ is defined as follows: consider the labelled directed graph $\mathcal{G}_i^x$ with node set $\mathcal{V} \cup \{\star\}$ and labelled edge set given by:

$$\begin{aligned}
&\bullet\ H_\ell \xrightarrow{k_{\ell\ell'}} H_{\ell'} && \text{if } k_{\ell\ell'} \neq 0 \text{ and } \ell \neq \ell' \\
&\bullet\ H_\ell \xrightarrow{\sum_{j\in W} k_{\ell j}} \star && \text{if } \sum_{j\in W} k_{\ell j} \neq 0 \\
&\bullet\ \star \xrightarrow{\lambda_{i\ell}(x)} H_\ell && \text{if } \lambda_{i\ell}(x) \neq 0
\end{aligned} \tag{3.7}$$

We recall some notion from graph theory: let $\mathcal{G}$ be a labelled directed graph. A labelled spanning tree of $\mathcal{G}$ rooted at some node g is a labelled directed graph ζ satisfying the following conditions:

i) the set of nodes of ζ coincides with the set of nodes of $\mathcal{G}$;

ii) any directed edge of ζ is a directed edge of $\mathcal{G}$, and the labels are conserved;

iii) ζ contains no cycle;

iv) for any node $g' \neq g$, in ζ there exists a directed path from g' to g.

Let $\Theta_i^x(\cdot)$ be the set of labelled spanning trees of $\mathcal{G}_i^x$ rooted at the argument, and let $w(\cdot)$ be the product of the edge labels of the tree in the argument. Then, $\mu_{i\ell}(x)$ is defined as

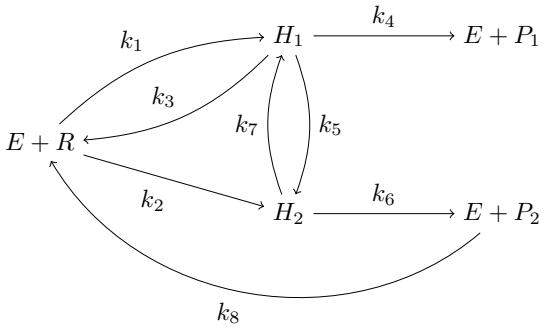
$$\mu_{i\ell}(x) = \frac{\sum_{\zeta \in \Theta_i^x(H_\ell)} w(\zeta)}{\sum_{\zeta \in \Theta_i^x(\star)} w(\zeta)}. \tag{3.8}$$

There might be no spanning tree rooted at a given intermediate species for some $x \in \mathbb{N}^{\mathcal{X}}$. In that case, $\mu_{i\ell}(x)$ is 0. The denominator is always strictly positive as any intermediate is eventually turned into a non-intermediate (Definition 2.1). The proof of the following result is given in Section 6.

Theorem 3.1. *For all $x \in \mathbb{N}^{\mathcal{X}}$, $i \in U$, $j \in W$, we have $\lambda_{ij}^*(x) = \widetilde{\lambda}_{ij}(x)$, hence (3.5) and (3.6) coincide.*

Below we give an example of a reduced SRN.

Example 3.1. Consider the reaction network with intermediate species H_1, H_2, taken with mass-action kinetics



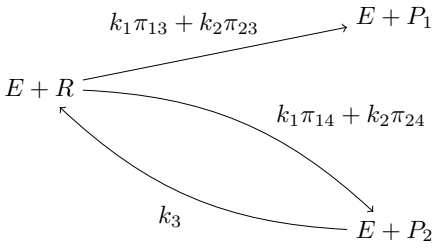
In this case there is only one initial reactant, namely $E+R$, while the final products are $E+R$, $E+P_1$ and $E+P_2$. Therefore the set of initial reactants and the set of final products have non-empty intersection. If we let $E+P_1 = y_3$ and $E+P_2 = y_4$, then, by summing the probabilities of all possible paths from H_1 to $E+P_1$, we find that

$$\pi_{13} = \frac{k_4}{k_3+k_4+k_5} \sum_{n\in\mathbb{N}} \left(\frac{k_5}{k_3+k_4+k_5} \cdot \frac{k_7}{k_6+k_7} \right)^n = \frac{k_4(k_6+k_7)}{(k_3+k_4)(k_6+k_7)+k_5k_6}.$$

Similarly, we calculate π_{14}, π_{23} and π_{24} and obtain

$$\begin{aligned}
\pi_{14} &= \frac{k_5k_6}{(k_3+k_4)(k_6+k_7)+k_5k_6}, \\
\pi_{23} &= \frac{k_4k_7}{(k_3+k_4)(k_6+k_7)+k_5k_6}, \\
\pi_{24} &= \frac{(k_3+k_4+k_5)k_6}{(k_3+k_4)(k_6+k_7)+k_5k_6}.
\end{aligned}$$

The reduced reaction network with mass-action kinetics is therefore



(3.9)

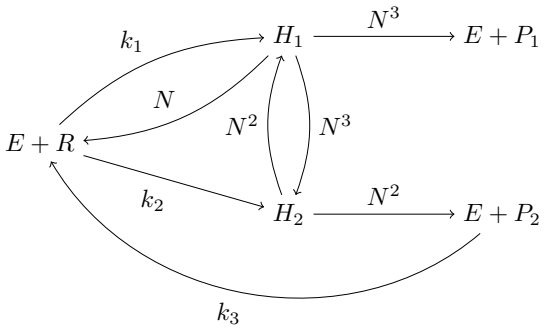
4 Results

Before formalising the setting and the assumptions, we provide some examples to motivate it. Recall Example 3.1. Intuitively, the reduced SRN behaves similarly to the original SRN if the time spent in intermediate states (states with at least one intermediate molecule being present) is insignificant compared to the time spent in other states. Thus, it is natural to consider situations for which the reaction rates out of intermediate states are all high, though this is not what is required for our results to hold (Example 4.6).

Consider a reaction network $(\mathcal{X}, \mathcal{C}, \mathcal{R})$ and a sequence of kinetics $\mathcal{K}^N$ indexed by $N \in \mathbb{N}$. Let $X_\cdot^N$ be the process (2.3) associated with the kinetics $\mathcal{K}^N$. Generally, we will have in mind that the consumption rates of the intermediates species increase in N. We will consider a multiscale setting, where the species abundances also are scaled according to N. Hence, we consider the asymptotic behaviour of the process $X_\cdot^N$ as $N \to \infty$, when both species abundances and rate constants depend on N, similarly to what is done in Ball et al. [2006], Kang and Kurtz [2013], Pfaffelhuber and Popovic [2013].

To increase readability, in the examples the reaction rates depending on N are simple powers of N with no prefactors (e.g. N^2 rather than kN^2). In the results these restrictions are not assumed and more general forms of reaction rates are allowed.

Example 4.1. Consider the SRN from Example 3.1 with rate constants



(4.1)

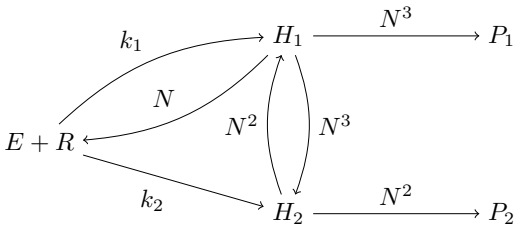
The reduced SRN has reaction rates given by (3.9) with

$$\pi_{13}^N = \frac{2N^3}{3N^3 + 2N}, \quad \pi_{14}^N = \frac{N^3}{3N^3 + 2N}, \quad \pi_{23}^N = \frac{N^3}{3N^3 + 2N}, \quad \pi_{24}^N = \frac{2N^3 + N}{3N^3 + 2N}.$$

We assume that the molecular abundances of R, P_1, P_2 are of order $O(N)$, while $X_t^N(E) = O(1)$. We further assume that at time 0 there are no intermediates present, that is, $X_0(H_1) = X_0(H_2) = 0$. The expression $O(N)$ will be made precise later, but it indicates that at a typical time $t > 0$, the molecular abundances of R, P_1, P_2 are of the same order of magnitude as N. With the assumption on the abundances, the rates of the reactions $E + R \to H_1$, $E + R \to H_2$ and $E + P_2 \to E + R$ are of order $O(N)$, while the intermediate species are consumed considerably faster. Therefore it seems reasonable that the intermediates might be eliminated from the description of the system and the dynamics described by the simpler reduced SRN in (3.9). We will show that the dynamics of the reduced SRN approximates the dynamics of (4.1) for N large. Specifically, we will show that the difference between the two stochastic processes associated with the two networks converges to 0 in the sense of Theorems 4.3 and 4.7 for $N \to \infty$.

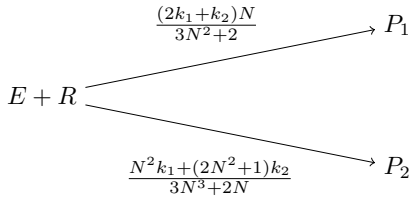
Example 4.2 (trapped in the intermediate chain). Consider the same reaction network as in Example 4.1, but with slightly changed reaction rates. The reaction $H_2 \to E + P_2$ is slowed down and has rate N (before N^2). The reaction $H_1 \to H_2$ is accelerated and has rate N^4 (before N^3). All other rates are left unchanged. We assume as before that the molecular abundances of R, P_1, P_2 are of order $O(N)$, while $X_t^N(E) = O(1)$. Although the intermediate species are consumed faster than the other species (the life time of a molecule of H_1 and of H_2 are of order $O(1/N^4)$ and $O(1/N^2)$, respectively), it is not possible to approximate the above SRN with one of the form (3.9), for any choice of kinetics. Indeed, it is more likely that an intermediate molecule is transformed into another intermediate molecule than into one of the two final products, $E + P_1$ and $E + P_2$. On average, an intermediate molecule will undergo the cycle of transformations $H_1 \to H_2 \to H_1$ N times before producing a non-intermediate complex. Since the life time of a molecule of H_2 is of order $O(1/N^2)$, the expected time until consumption of such a cycle of intermediates is of the order $O(1/N)$, while the rate of production of intermediate molecules is of order $O(N)$ when molecules of E are present, according to the hypothesis $X_t^N(R) = O(N)$. This will result in a positive number of intermediate species being present at any fixed time t. Therefore, in this case, the intermediate species cannot be eliminated in the sense of this paper.

Example 4.3 (rescaling of time). Consider the following SRN, which is a modified version of (4.1). The enzyme E is removed from the product complexes $E + P_1$ and $E + P_2$, and the reaction $E + P_2 \to E + R$ is deleted:



Assume that the molecular abundance of R is of order $O(N)$ and that the molecular abundance of E is of order $O(1)$. The small amount of enzyme molecules will be consumed fast and none will be produced. Therefore, after a while, there will be no enzyme molecules present. Each intermediate molecule will fast produce P_1 or P_2 and, after that, no other reaction can possibly take place. That is, after a time of order $O(1/N)$, no reaction will take place. Thus, in order to observe the dynamics of the system, time should be rescaled by a factor N. That is, the time $\tilde{t} = t/N$ should be considered. This is the same as studying the SRN with all reaction rates rescaled by a factor of $1/N$.

Despite some reaction rates tend to zero with N, our results can be applied to approximate the dynamics of the SRN. In particular the reduced SRN is given by



where the magnitudes of the molecular abundances of E, R, P_1, P_2 are the same as in the full reaction network.

4.1 Assumptions

Let $(\mathcal{X}, \mathcal{C}, \mathcal{R})$ be a SRN with a set of intermediate species $\mathcal{V} \subset \mathcal{X}$, let $\mathcal{K}^N$ be a sequence of kinetics indexed by $N \in \mathbb{N}$, and let $X_\cdot^N$ be the corresponding stochastic process (2.3). Define

$$\mathcal{R}^0 = \{y_i \to y_j \in \mathcal{R} \colon y_i \notin \mathcal{V}\}, \tag{4.2}$$

$$\mathcal{R}^1 = \{y_i \to y_j \in \mathcal{R} \colon y_i, y_j \notin \mathcal{V}\} \subset \mathcal{R}^0. \tag{4.3}$$

Specifically, $\mathcal{R}^0$ is the set of reactions whose reactant is not an intermediate, while $\mathcal{R}^1$ is the set of reactions not involving intermediates at all.

Fix a non-negative vector of scaling coefficients, $\alpha = (\alpha(S))_{S \in \mathcal{X} \setminus \mathcal{V}} \in \mathbb{R}_{\geq 0}^{\mathcal{X} \setminus \mathcal{V}}$, and define the rescaled process,

$$\widehat{X}_t^N = N^{-\alpha} p(X_t^N), \tag{4.4}$$

where $p \colon \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X} \setminus \mathcal{V}}$ is the projection onto the non-intermediate species space and the multiplication $N^{-\alpha} p(X_t^N)$ is intended component-wise. The process $\widehat{X}_\cdot^N$ is the rescaled process in the sense of Ball et al. [2006], Kang and Kurtz [2013], Pfaffelhuber and Popovic [2013] for the non-intermediate species. Since $\alpha(S)$ might differ from species to species, $\widehat{X}_\cdot^N$ is a multiscale process.

Assumption 2. Let α be given as in (4.4).

(i) (Rate functions and intermediates) We assume that $(\mathcal{X}, \mathcal{C}, \mathcal{R})$ equipped with $\mathcal{K}^N$ satisfies Assumption 1 for all $N \in \mathbb{N}$.

(ii) (Rescaling of abundances) We assume that for any non-intermediate species $S \in \mathcal{X} \setminus \mathcal{V}$,

$$\widehat{X}_t^N(S) = \mathcal{O}(1), \tag{4.5}$$

that is, the scaled abundances do not blow up before time t. To make (4.5) precise, we require that there exists $T > 0$ such that for any $S \in \mathcal{X} \setminus \mathcal{V}$,

$$\begin{cases} \forall \nu > 0 \; \exists \Upsilon_\nu \colon \limsup\limits_{N \to \infty} P\left(\sup\limits_{[0,T]} \widehat{X}_t^N(S) > \Upsilon_\nu \right) < \nu & \text{(4.6a)} \\ \mathcal{L}\left\{ t \in [0,T] \colon \lim\limits_{N \to \infty} \widehat{X}_t^N(S) = 0 \text{ a.s.} \right\} = 0, & \text{(4.6b)} \end{cases}$$

where $\mathcal{L}$ denotes the usual Lebesgue measure on $\mathbb{R}$.

(iii) (Convergence of rate functions) We assume that there exist a set of locally Lipschitz functions $\{\lambda_r(\cdot)\}_{r\in\mathcal{R}^0}$ defined on $\mathbb{R}_{\geq 0}^{\mathcal{X}\setminus\mathcal{V}}$, fulfilling

$$x \in \mathbb{R}_{>0}^{\mathcal{X}\setminus\mathcal{V}} \Rightarrow \lambda_r(x) > 0,$$

and a set of non-negative real numbers $\{\beta_r\}_{r\in\mathcal{R}^0}$ such that, for all $r \in \mathcal{R}^0$,

$$N^{-\beta_r}\lambda_r^N(N^\alpha x) \xrightarrow[N\to\infty]{} \lambda_r(x) \tag{4.7}$$

uniformly on compact sets, where the rate functions λ_r^N are extended to the real vectors by considering the floor function of the argument.

(iv) (Degradation of intermediates) Let C_n^N, τ_n^N, t_n^N and $\pi_{\ell j}^N$ be as defined after Assumption 1. Let

$$\beta_\ell^* = \max_{i\in U}\beta_{i\ell}, \qquad \alpha_j^* = \min_{S\in y_j}\alpha(S),$$

where $\beta_{i\ell}$ is as in (iii) for $r = y_i \to H_\ell$. Moreover, define

$$p_{\ell j}^\varepsilon(N) = P\left(\tau_1^N > \frac{N^{\alpha_j^*}\varepsilon}{N^{\beta_\ell^*}\pi_{\ell j}^N}\middle| C_1^N(t_1^N) = H_\ell, C_1^N(t_1^N+\tau_1^N) = y_j\right). \tag{4.8}$$

By definition of the continuous-time Markov chains $C_n^N(\cdot)$, for any n

$$P\left(\tau_n^N > \frac{N^{\alpha_j^*}\varepsilon}{N^{\beta_\ell^*}\pi_{\ell j}^N}\middle| C_n^N(t_n^N) = H_\ell, C_n^N(t_n^N+\tau_n^N) = y_j\right) = p_{\ell j}^\varepsilon(N).$$

We assume that the size of τ_n^N is controlled, that is, for all $\varepsilon > 0$, $\ell \in \bigcup_{i\in U} V_i$ and $j \in W$, we have

$$\pi_{\ell j}^N N^{\beta_\ell^*-\alpha_j^*} p_{\ell j}^\varepsilon(N) \xrightarrow[N\to\infty]{} 0. \tag{4.9}$$

Sufficient conditions for (4.9) are given in Propositions 4.1 and 4.2.

(v) (Single scale system) For any non-intermediate species $S \in \mathcal{X}\setminus\mathcal{V}$, let

$$\mathcal{R}_S^1 = \mathcal{R}_S \cap \mathcal{R}^1 \qquad \text{and} \qquad \overline{\mathcal{R}}_S = \left\{r \in \mathcal{R}^* \setminus \mathcal{R}^1 :\ \xi_r(S) \neq 0\right\}.$$

Moreover, for all $\ell \in V_i$ and j in the set of complexes indices, let $\pi_{\ell j}^N$ be as in (3.3). We assume that

$$\begin{cases} \exists\, \gamma_{\ell j} = \lim\limits_{N\to\infty} \log_N \pi_{\ell j}^N & \in [-\infty, 0] \\ \exists \lim\limits_{N\to\infty} \pi_{\ell j}^N N^{-\gamma_{\ell j}} & \text{if } \gamma_{\ell j} > -\infty. \end{cases} \tag{4.10}$$

and

$$\max\left(\{\beta_r\}_{r\in\mathcal{R}_S^1} \cup \{\beta_{i\ell} + \gamma_{\ell j}\}_{\ell\in V_i, y_i\to y_j\in\overline{\mathcal{R}}_S}\right) \leq \alpha(S), \tag{4.11}$$

where β_r with $r \in \mathcal{R}^0$ is as in (iii), and $\max\emptyset = -\infty$.

Remark 4.1. 'Single scale system' in Assumption 2(v) refers to the time scale of the reduced SRN, as defined in Pfaffelhuber and Popovic [2013].

Remark 4.2. Time rescaling in the sense of Example 4.3 might be considered. It is equivalent to a rescaling of all the rate functions by a common factor, and therefore equivalent to adding a common term to all the β's . Thus, time rescaling is implicitly considered in our framework of model reduction. We will ignore it in the development of the theory.

Remark 4.3. Assume mass-action kinetics and assume that for any reaction $r\colon y_i \to y_j \in \mathcal{R}^0$, the constant k_r^N is of the form $N^{\eta_r}k_r$ with $k_r > 0$ and $\eta_r \in \mathbb{R}$. Thus,

$$\lambda_r^N(N^\alpha x) = N^{\eta_r}k_r \prod_{S\in y_i} \frac{(N^{\alpha(S)}x(S))!}{(N^{\alpha(S)}x(S)-y_i(S))!}\mathbb{1}_{\{N^{\alpha(S)}x(S)\geq y_i(S)\}}\,.$$

This means that the right scaling for the rate function λ_r^N is

$$\beta_r = \eta_r + \sum_{S\in y_i}\alpha(S)\cdot y_i(S)\,.$$

Indeed,

$$N^{-\beta_r}\lambda_r^N(N^\alpha x) \xrightarrow[N\to\infty]{} \lambda_r(x)$$

uniformly on compact sets, where

$$\lambda_r(x) = k_r \prod_{\substack{S\in y_i\\ \alpha(S)=0}} \frac{x(S)!}{(x(S)-y_i(S))!}\mathbb{1}_{\{x(S)\geq y_i(S)\}} \prod_{\substack{S\in y_i\\ \alpha(S)>0}} x(S)^{y_i(S)}\mathbb{1}_{\{x(S)>0\}}\,.$$

Remark 4.4. Theorems 4.3 and 4.5 below hold even if (4.10) and (4.11) in Assumption 2(v) are replaced by the weaker conditions

$$\exists c_{\ell j} > 0 \text{ s.t. } \limsup_{N\to\infty} \pi_{\ell j}^N N^{\beta_\ell^* - \alpha_j^*} \leq c_{\ell j} \tag{4.12}$$

$$\max\left(\{\beta_r\}_{r\in\mathcal{R}_S^1} \cup \left\{\limsup_{N\to\infty}\left(\beta_{i\ell} + \log_N \pi_{\ell j}^N\right)\right\}_{\ell\in V_i, y_i\to y_j\in\overline{\mathcal{R}}_S}\right) \leq \alpha(S). \tag{4.13}$$

We will use these in the proof of Theorems 4.3 and 4.5.

Under the assumption that $\widehat{X}_0^N$ is bounded uniformely on N, condition (4.6a) is fulfilled for a special class of reaction networks called *conservative* reaction networks (cf. Remark 4.6). In order to state suffiecient conditions for (4.9) to hold, for any $\ell \in V_i$ we define

$$a_\ell = \min_{y_j\in\mathcal{W}_\ell} \alpha_j^*,$$

where $\mathcal{W}_\ell \subseteq \mathcal{W}$ denotes the set of final products which are obtainable from H_ℓ through a path of intermediates. In other words, $\mathcal{W}_\ell$ is the set of final products y_j such that there exists a path of the form

$$H_\ell \to H_{\ell_1} \to \cdots \to H_{\ell_k} \to y_j.$$

The following holds:

Proposition 4.1. *Equation (4.9) holds if for all $\ell \in \bigcup_{i\in U} V_i$ and $\varepsilon > 0$, we have*

$$N^{\beta_\ell^* - a_\ell} P\left(\tau_1^N > N^{a_\ell - \beta_\ell^*}\varepsilon \,\middle|\, C_1^N(t_1^N) = H_\ell\right) \xrightarrow[N\to\infty]{} 0. \tag{4.14}$$

Moreover, (4.9) holds if for all $\ell \in \bigcup_{i\in U} V_i$ and $\varepsilon > 0$, we have (4.12) and

$$N^{\beta_\ell^* - a_\ell} E\left[\tau_1^N \,\middle|\, C_1^N(t_1^N) = H_\ell\right] \xrightarrow[N\to\infty]{} 0. \tag{4.15}$$

Proof. The first part of the proposition is proven by

$$\begin{aligned}\sum_{j\in W} \pi_{\ell j}^N N^{\beta_\ell^* - \alpha_j^*} p_{\ell j}^\varepsilon(N) &\leq N^{\beta_\ell^* - a_\ell} \sum_{j\in W} \pi_{\ell j}^N P\left(\tau_1^N > N^{a_\ell - \beta_\ell^*}\varepsilon \,\middle|\, C_1^N(t_1^N) = H_\ell, C_1^N(t_1^N + \tau_1^N) = y_j\right)\\ &= N^{\beta_\ell^* - a_\ell} P\left(\tau_1^N > N^{a_\ell - \beta_\ell^*}\varepsilon \,\middle|\, C_1^N(t_1^N) = H_\ell\right).\end{aligned}$$

The second part of the proposition follows from

$$\sum_{j\in W} \pi^N_{\ell j} N^{\beta^*_\ell - \alpha^*_j} E\left[\tau^N_1 \middle| C^N_1(t^N_1) = H_\ell, C^N_1(t^N_1+\tau^N_1) = y_j\right]$$
$$\leq N^{\beta^*_\ell - a_\ell} \sum_{j\in W} \pi^N_{\ell j} E\left[\tau^N_1 \middle| C^N_1(t^N_1) = H_\ell, C^N_1(t^N_1+\tau^N_1) = y_j\right]$$
$$= N^{\beta^*_\ell - a_\ell} E\left[\tau^N_1 \middle| C^N_1(t^N_1) = H_\ell\right].$$

Therefore, (4.15) implies that for any $j \in W$

$$\pi^N_{\ell j} N^{\beta^*_\ell - \alpha^*_j} E\left[\tau^N_1 \middle| C^N_1(t^N_1) = H_\ell, C^N_1(t^N_1+\tau^N_1) = y_j\right] \xrightarrow[N\to\infty]{} 0.$$

By Markov inequality, this implies that $p^\varepsilon_{\ell j}(N)$ tends to zero as N goes to infinity. By (4.12), the latter leads to (4.9), and the proof is complete. □

Since τ^N_n is a phase-type distributed random variable, we can express (4.14) in terms of the exponential of the transition rate matrix (3.2). Specifically, (4.14) is equivalent to

$$N^{\beta^*_\ell - a_\ell} (e_\ell)^\top \exp\left(N^{a_\ell - \beta^*_\ell} \varepsilon Q^N_{V,V}\right) e \xrightarrow[N\to\infty]{} 0 \qquad \forall \ell \in \bigcup_{i\in U} V_i,$$

where $(e_\ell)^\top$ denotes the transpose of the canonical base vector with a one in the ℓ-th entry and e is the vector with all entries equal to one. A sufficient condition for (4.15) to hold is given in the proposition below:

Proposition 4.2. *Assume Assumptions 2(i,iii) are fulfilled for some $\alpha \in \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$. For all $i \in U$ and $\ell \in V$, let $\mu^N_{i\ell}(x)$ be as in (3.8) and define*

$$\alpha^* = \min_{j\in W} \alpha^*_j.$$

We have that, if

$$N^{-\alpha^*} \mu^N_{i\ell}(N^\alpha x) \xrightarrow[N\to\infty]{} 0 \tag{4.16}$$

for all $x \in \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}_{\geq 0}$ and for all $i \in U$, $\ell \in V$, then (4.15) in Assumption 2(iv) holds. Moreover, if $\alpha^*_j = \alpha^*_{j'}$ for all $j, j' \in W$, then (4.16) is also a necessary condition for (4.15) to hold.

We prove Proposition 4.2 in Section 6. The condition (4.15) is sufficient for (4.9) to hold, but it is not necessary, as shown in Example 4.6. Before moving on, we make a number of remarks.

4.2 The Process $Z^N_\cdot$

In order to show that the reduced SRN provides a good approximation, under the given assumptions, of different features of the original SRN, we define a sequence of processes $Z^N_\cdot$ ad hoc. We choose them such that for any fixed t the (rescaled) difference $\left|X^N_t - Z^N_t\right|$ tends to zero in probability, and such that the process $Z^N_\cdot$ is distributed as the process associated with the reduced SRN. We will prove other convergence statements in Theorems 4.3, 4.5 and 4.7.

Recall that $p\colon \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$ is the projection onto the non-intermediate species space. By Assumption 2(i), the reaction rates $\lambda^N_r(\cdot)$ with $r \in \mathcal{R}^0$ do not depend on the counts of intermediates. That is, for any $x \in \mathbb{N}^{\mathcal{X}}$,

$$\lambda^N_r(x) = \bar{\lambda}^N_r(p(x)),$$

for some function $\bar{\lambda}^N_r\colon \mathbb{N}^{\mathcal{X}\setminus\mathcal{V}} \to \mathbb{R}_{\geq 0}$. For the sake of convenience, we will abuse notation and let $\bar{\lambda}^N_r(x) = \lambda^N_r(x)$ for all $x \in \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$.

Given the n-th chain of intermediates $C^N_n(\cdot)$ appearing in relation to the process $X^N_\cdot$, we denote by $\{C^N_n(\cdot) \in C_{i\ell j}\}$ the event that $C^N_n(\cdot)$ originates from the reaction $y_i \to H_\ell$ and eventually produces the final complex y_j. Such an event is measurable with respect to the σ-algebra σ^N_∞ as introduced in

(3.4). Furthermore, let $M^N_{i\ell j}(t)$ denote the number of the chains originated before time t and such that $\{C^N_n(\cdot) \in C_{i\ell j}\}$:

$$M^N_{i\ell j}(t) = \#\left\{n\colon C^N_n(\cdot) \in C_{i\ell j}\,,\, t^N_n \le t\right\} = \sum_{n=1}^{Y_{i\ell}\left(\int_0^t \lambda^N_{i\ell}(X^N_s)ds\right)} \mathbb{1}_{\{C^N_n(\cdot)\in C_{i\ell j}\}}.$$

The processes $M^N_{i\ell j}(\cdot)$ are therefore arrival processes, and we might represent them in terms of independent and identically distributed unit-rate Poisson processes $Y_{i\ell j}(\cdot)$ such that

$$M^N_{i\ell j}(t) = Y_{i\ell j}\left(\int_0^t \pi^N_{\ell j}\lambda^N_{i\ell}(X^N_s)ds\right). \tag{4.17}$$

In this context, $Y_{i\ell}(t) = \sum_{j\in W} Y_{i\ell j}(t)$. Moreover, let $t^N_{i\ell j,n}$ be the time of the n-th jump of the process $M^N_{i\ell j}(\cdot)$, and let $\tau^N_{i\ell j,n}$ be a collection of independent random variables distributed as τ^N_1 given $(C^N_1 \in C_{i\ell j})$. We now consider the process counting the number of chains of intermediates $C^N_n(\cdot)$ consumed before time t and such that $\{C^N_n(\cdot) \in C_{i\ell j}\}$. Such a process is distributed as

$$\overline{M}^N_{i\ell j}(t) = \sum_{n=1}^{M^N_{i\ell j}(t)} \mathbb{1}_{\{t^N_{i\ell j,n}+\tau^N_{i\ell j,n}\le t\}}.$$

For any time t, we have $M^N_{i\ell j}(t) \ge \overline{M}^N_{i\ell j}(t)$. The process $\widehat{X}^N_\cdot$ can be equivalently expressed as

$$\widehat{X}^N_t = \widehat{X}^N_0 + N^{-\alpha}\left[\sum_{r\in\mathcal{R}^1}\xi_r Y_r\left(\int_0^t \lambda^N_r(X^N_s)ds\right) + \sum_{i\in U}\sum_{j\in W}\left(y_j\sum_{\ell\in V_i}\overline{M}^N_{i\ell j}(t) - y_i\sum_{\ell\in V_i}M^N_{i\ell j}(t)\right)\right], \tag{4.18}$$

where the Poisson processes $Y_r(\cdot)$ are the same as those appearing in (2.3). We will use this representation in the remaining part of the paper.

We define the process $\widehat{Z}^N_\cdot$ on $N^{-\alpha}\mathbb{N}^{\mathcal{X}\setminus\mathcal{V}}$ as

$$\widehat{Z}^N_t = \widehat{Z}^N_0 + N^{-\alpha}\left[\sum_{r\in\mathcal{R}^1}\xi_r Y_r\left(\int_0^t \lambda^N_r(Z^N_s)ds\right) + \sum_{i\in U}\sum_{j\in W}(y_j-y_i)\sum_{\ell\in V_i}Y_{i\ell j}\left(\int_0^t \pi^N_{\ell j}\lambda^N_{i\ell}(Z^N_s)ds\right)\right]. \tag{4.19}$$

For any fixed $t \ge 0$, the random variables $\widehat{X}^N_t$ and $\widehat{Z}^N_t$ are measurable with respect to

$$\sigma\left(Y_r(s), Y_{i\ell j}(s), \tau^N_{i\ell j,n}\colon r\in\mathcal{R}^1, i\in U, \ell\in V_i, j\in W, n, N\in\mathbb{N} \text{ and } 0\le s<\infty\right).$$

The above σ-algebra contains information about the Poisson processes $Y_r(\cdot)$ for reactions not involving intermediates, about the Poisson processes $Y_{i\ell j}(\cdot)$ that drive $M^N_{i\ell j}(\cdot)$ and about the *delays* $\tau^N_{i\ell j,n}$ of the reactions proceeding through intermediates species. It does not contain full information on the intermediate chains $C^N_n(\cdot)$, but that is not required in the description of the processes $\widehat{X}^N_\cdot$ and $\widehat{Z}^N_\cdot$. The random variables we are interested in will all be measurable with respect to the above σ-algebra, and therefore are defined on the same probability space. Since $\widehat{Z}^N_t$ is, up to rescaling, expressed in the form (2.3), it is distributed as the rescaled stochastic process associated with (3.1).

There is a precise intuition behind the choice of $\widehat{Z}^N_t$ as approximating process for the original system. Consider (4.18): if (4.9) holds, then we expect the lifetime of the intermediate species to decrease with N. Thus, we could imagine that, for any fixed time t, $M^N_{i\ell j}(t) = \overline{M}^N_{i\ell j}(t)$ with high probability and, thus, that $\widehat{X}^N_t$ is approximated by

$$\widehat{W}^N_t = \widehat{X}^N_0 + N^{-\alpha}\left(\sum_{r\in\mathcal{R}^1}\xi_r Y_r\left(\int_0^t \lambda^N_r(X^N_s)ds\right) + \sum_{i\in U}\sum_{j\in W}(y_j-y_i)\sum_{\ell\in V_i}M^N_{i\ell j}(t)\right). \tag{4.20}$$

The process $\widehat{Z}^N_\cdot$ in (4.19) is defined analogously to (4.20).

Unfortunately, we cannot hope for $\widehat{X}^N_\cdot$ to converge weakly to $\widehat{Z}^N_\cdot$ in the Skorohod topology in general (cf. Example 5.3). However, we will show a uniform punctual convergence in probability as well as convergence in occupation measure for the difference of the stopped processes $\widehat{X}^N_{\cdot\wedge T}$ and $\widehat{Z}^N_{\cdot\wedge T}$, for any fixed $T > 0$. Furthermore, we will give additional hypothesis under which the convergence in probability in the Skorohod space holds.

4.3 Bounded Reaction Rates

Recall that $\mathcal{R}^0$ in (4.2) is the set of reactions whose reactant is not an intermediate. Here we are concerned with the case when all reaction rates of reactions in $\mathcal{R}^0$ are bounded by a power of N, specifically for any $r \in \mathcal{R}^0$,

$$N^{-\beta_r}\lambda_r^N(x) \leq B_r \quad \forall\, N \in \mathbb{N},\, \forall\, x \in \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}_{\geq 0}, \tag{4.21}$$

where β_r is as in Assumption 2(iii) and B_r is a positive constant (later the constant will also be referred to as $B_{i\ell}$ if in relation to the reaction $y_i \to H_\ell$). It is worth mentioning that in this case, (4.6a) in Assumption 2(ii) is always fulfilled if $\widehat{X}^N_0$ is stochastically bounded (cf. Remark 4.5). This is desirable because it suffices to control stochastic boundeness of a real random variable rather than of an entire stochastic process. Moreover, (4.21) can be assume to hold if the network is conservative and $\widehat{X}^N_0$ is bounded independently of N (cf. Remark 4.6).

The proofs of Theorems 4.3 and 4.5 can be found in Section 7, using the relaxed version of Assumption 2(v) as given in Remark 4.4. The weaker condition is sufficient to prove Corollary 4.4 as well.

Theorem 4.3. *Assume Assumption 2 is fulfilled for some $\alpha \in \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$. Further, assume that*

$$E\left[\left|\widehat{X}^N_0 - \widehat{Z}^N_0\right|\right] \xrightarrow[N\to\infty]{} 0,$$

and that the initial amounts of the intermediate species are 0. Finally, assume that for any $r \in \mathcal{R}^0$, (4.21) holds and λ_r is Lipschitz. Then, if T is as in Assumption 2(ii), we have that

$$\sup_{t\in[0,T]} E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right] \xrightarrow[N\to\infty]{} 0, \tag{4.22}$$

In particular, (4.22) implies that for all $\varepsilon > 0$,

$$\sup_{t\in[0,T]} P\left(\left|\widehat{X}^N_t - \widehat{Z}^N_t\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0. \tag{4.23}$$

Finally, for any continuous function $f\colon \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}} \to \mathbb{R}$ we have

$$P\left(\sup_{t\in[0,T]}\left|\int_0^t \left(f(\widehat{X}^N_s) - f(\widehat{Z}^N_s)\right) ds\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0. \tag{4.24}$$

Remark 4.5. Assume that (4.21), (4.12), and (4.13) hold. Assume further that $\widehat{X}^N_0$ is stochastically bounded, meaning that for every $\nu > 0$ there exists Υ_ν such that for every $S \in \mathcal{X}\setminus\mathcal{V}$

$$\limsup_{N\to\infty} P\left(\widehat{X}^N_0(S) > \Upsilon_\nu\right) < \nu.$$

Our aim is to prove (4.6a). By (4.18)

$$\sup_{t\in[0,T]} \widehat{X}^N_t(S) \leq X^N_0(S) + N^{-\alpha(S)} \sum_{r\in\mathcal{R}^1_S} |\xi_r(S)|\, Y_r(N^{\beta_r}B_r T) + \\ + N^{-\alpha(S)} \sum_{i\in U}\sum_{j\in W} 2\left(y_j(S) + y_i(S)\right) \sum_{\ell\in V_i} Y_{i\ell j}(\pi^N_{\ell j} N^{\beta_{i\ell}} B_{i\ell} T),$$

where $\mathcal{R}^1_S$ is defined according to (2.2). Using assumptions (4.12), (4.13) and the Law of Large Numbers for Poisson processes to control the above expression for $\alpha(S) > 0$, we obtain that, for any $\nu > 0$, there exists $\Upsilon'_\nu > 0$, such that

$$\limsup_{N\to\infty} P\left(\sup_{t\in[0,T]} \widehat{X}^N_t(S) > \Upsilon'_\nu\right) < \nu.$$

Remark 4.6. Conservative reaction networks are a special class of reaction networks [Horn and Jackson, 1972]. In a conservative reaction network, a positive linear combination of the species abundances is preserved throughout time and, hence, the total abundances are bounded from above given any initial condition. In such class of reaction networks, if $\widehat{X}_0^N$ is bounded uniformly on N then condition (4.21) is fulfilled. Indeed, if the original reaction network is conservative, then the reduced reaction network is conservative as well [Feliu and Wiuf, 2013a]. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ denote the spaces spanned by the reaction vectors of the original and of the reduced network, respectively. Moreover let $\mathcal{S} = p(\mathcal{S}_1) \cup \mathcal{S}_2 \subset \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$. It can be shown that $\mathcal{S}_2 \subseteq p(\mathcal{S}_1)$, but this lies outside our concerns. The initial condition $\widehat{X}_0^N$ varies in a compact set K_0. Therefore, for any $r \in \mathcal{R}^0$, we might consider a modified version of the rate functions λ_r^N, such that

$$N^{-\beta_r}\lambda_r^N(N^\alpha x) = 1 \quad \forall x \notin (S + K_1) \cap \mathbb{R}_{\geq 0}^{\mathcal{X}\setminus\mathcal{V}},$$

and $K_1 \supset K_0$ is a compact set. Thus, the limit functions λ_r in Assumption 2(iii) are 1 outside a compact set and therefore bounded. Due to (4.7), condition (4.21) is met. In particular, it follows from Remark 4.5 that in this case (4.6a) always holds.

Corollary 4.4. *Assume that the assumptions of Theorem 4.3 hold. Then, the difference between the processes $\widehat{X}^N_{\cdot\wedge T}$ and $\widehat{Z}^N_{\cdot\wedge T}$ converges in finite dimensional distribution to 0.*

Proof. From Theorem 4.3 we have that (4.23) holds for any $\varepsilon > 0$. Thus, for any finite set of time points $\{t_m\}_{m=0}^p \subseteq [0, T]$ we have that

$$\begin{aligned} P\left(\max_{0\leq m\leq p}\left|\widehat{X}^N_{t_m} - \widehat{Z}^N_{t_m}\right| > \varepsilon\right) &= P\left(\bigcup_{m=0}^{p}\left\{\left|\widehat{X}^N_{t_m} - \widehat{Z}^N_{t_m}\right| > \varepsilon\right\}\right) \\ &\leq \sum_{m=0}^{p} P\left(\left|\widehat{X}^N_{t_m} - \widehat{Z}^N_{t_m}\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0, \end{aligned}$$

hence the corollary holds. □

We discuss here some applications of Theorem 4.3 and Corollary 4.4.

Example 4.4. Consider the reaction network in Example 4.1. Assumption 2(i) holds. Further, if we let $\alpha(E) = 0$ and $0 < \alpha(R) = \alpha(P_1) = \alpha(P_2) < 2$, then Assumption 2(ii-v) are satisfied if we choose the initial value X_0^N proportional to the scaling N^α and β_r according to Remark 4.3. Note that the reaction network is conservative in the sense of Remark 4.6. Thus, (4.21) holds and by Theorem 4.3 and Corollary 4.4, the probability distribution of the process associated with the reduced SRN approximates, in the sense of Theorem 4.3 and Corollary 4.4, the probability distribution of the process (4.1).

Example 4.5. Consider the Michaelis-Menten mechanism taken with mass-action kinetics:

$$E + R \underset{k_1 N^{\eta_1}}{\overset{k_0}{\rightleftarrows}} H \xrightarrow{k_2 N^{\eta_2}} E + P$$

Assumption 2(i) is satisfied, as well as (4.21) since the network is conservative. The probability that a molecule of H is transformed into the complex $E + R$ is $k_1 N^{\eta_1} / (k_1 N^{\eta_1} + k_2 N^{\eta_2})$, while the probability that it is transformed into the complex $E + P$ is $k_2 N^{\eta_2} / (k_1 N^{\eta_1} + k_2 N^{\eta_2})$. The reduced SRN is given by

$$E + R \xrightarrow{\frac{k_0 k_2 N^{\eta_2}}{k_1 N^{\eta_1} + k_2 N^{\eta_2}}} E + P$$

If we let that $\alpha(E) = 0$, $\alpha(R) < \eta_1 \vee \eta_2$ and $\alpha(P) = \alpha(R) \wedge (\alpha(R) + \eta_2 - \eta_1)$, then Assumption 2(ii-v) are satisfied if we choose the initial value X_0^N proportional to the scaling N^α and β_r according to Remark 4.3. In this case, Theorem 4.3 and Corollary 4.4 state in which sense the original process is approximated by the one associated with the reduced SRN. The magnitudes of the molecular abundances are the same as in the original system.

In the reduced SRN the amount of enzyme E is conserved. Hence, the model can further be reduced to

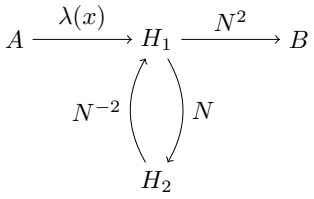
$$R \xrightarrow{\frac{E^0 k_0 k_2 N^{\eta_2}}{k_1 N^{\eta_1} + k_2 N^{\eta_2}}} P$$

where the amount of E molecules constantly equals E^0.

Let $\delta = \alpha(R) + \min\{0, \eta_2 - \eta_1\}$. If $\delta < 0$, we wait a time of order $O(N^{-\delta})$ for the first reaction to occur in the reduced SRN. Thus, we might rescale time in the original SRN by $\widetilde{t} = N^{\delta} t$. As shown in example 4.3, this is equivalent to rescale the rate functions. After rescaling, reduction can be performed again to obtain an approximation of the system's dynamics.

The following example concerns a network where not all the rates out of intermediate states are high. Moreover, it shows that condition (4.15) is sufficient for (4.9) in Assumption 2(iv) to hold, but it is not necessary.

Example 4.6. Consider the SRN taken with mass-action kinetics,



with $\alpha(A) = \alpha(B) = 0$. Assumption 2 is fulfilled if we choose the initial value X_0^N proportional to the scaling N^{α} and β_r according to Remark 4.3. This is true even though the consumption rate of H_2 tends to zero. Moreover, the reaction network is conservative, thus by Theorem 4.3, the reduced SRN

$$A \xrightarrow{\lambda(x)} B$$

provides a good approximation of the dynamics of the original SRN, for N large.

Further, (4.14) holds since for any fixed $\varepsilon > 0$, the probability that a chain of intermediates survives for a time bigger than ε goes to zero with $N \to \infty$. Hence by Proposition 4.1 (4.9) holds as well. However, in this case (4.15) does not hold. If we denote $A = y_3$ and $B = y_4$, this can be shown by making use of Proposition 4.2 and

$$\mu_{32}^N(x) = \frac{Nk\lambda(x)}{N^2 \cdot N^{-2}} = Nk\lambda(x) \xrightarrow[N\to\infty]{} \infty \quad \text{for any } x \in \mathbb{R}_{\geq 0}^{\mathcal{X}\setminus\mathcal{V}}.$$

For the particular case $\alpha > 0$, a stronger convergence result than those stated in Theorem 4.3 holds. The result does not hold generally for all α, as shown in Example 5.3.

Theorem 4.5. *Assume the assumptions of Theorem 4.3 are fulfilled and that $\alpha > 0$. Then, for any $\varepsilon > 0$,*

$$P\left(\sup_{t\in[0,T]} \left|\widehat{X}_t^N - \widehat{Z}_t^N\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0. \tag{4.25}$$

In particular, this implies that the difference between the processes $\widehat{X}_{\cdot\wedge T}^N$ and $\widehat{Z}_{\cdot\wedge T}^N$ converges weakly to 0 in the Skorohod topology.

4.4 Unbounded Reaction Rates

In this section, we will relax the hypothesis of boundedness in Theorem 4.3. To begin with, we introduce some new notation. Assume Assumption 2 is fulfilled and let $\mathcal{R}^*$ be defined as in (3.1). Define

$$\beta_{ij}^* = \max_{\ell\in V_i} \{\beta_{ij}, \beta_{i\ell} + \gamma_{\ell j}\},$$

where $\beta_{ij}, \beta_{i\ell}$ is as in Assumption 2(iii). We have that for any reaction $r \in \mathcal{R}^*$,

$$N^{-\beta_r^*} \lambda_r^{N,*}(Z_t^N) \xrightarrow[N\to\infty]{} \lambda_r^*(\widehat{Z}_t), \tag{4.26}$$

where $\lambda_r^{N,*}(\cdot)$ is defined in (3.5) and $\{\lambda_r^*\}_{r\in\mathcal{R}^*}(\cdot)$ is a set of locally Lipschitz functions such that

$$v \in \mathbb{R}^{\mathcal{X}}_{>0} \Rightarrow \lambda_r^*(v) > 0$$

(Assumption 2(iii)). As in Pfaffelhuber and Popovic [2013], we distinguish between fast and slow reactions. Let

$$\mathcal{R}^f = \bigcup_{S\colon \alpha(S)>0} \{y_i \to y_j \in \mathcal{R}_S^*\colon \alpha(S) = \beta_{ij}^*\}$$
$$\mathcal{R}^s = \bigcup_{S\colon \alpha(S)=0} \{y_i \to y_j \in \mathcal{R}_S^*\colon \alpha(S) = \beta_{ij}^*\}.$$

Moreover, let the vector $\xi_r^* \in \mathbb{R}^{\mathcal{X}}$ be defined by its entries

$$\xi_r^*(S) = \lim_{N\to\infty} N^{\beta_r^* - \alpha(S)} \xi_r(S).$$

Specifically, $\xi_r^*(S) = \xi_r(S)$, if $\alpha(S) = \beta_r^*$, and $\xi_r^*(S) = 0$, otherwise.

Lemma 4.6. *Assume Assumption 2 is fulfilled for some $\alpha \in \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$ and let T be as in Assumption 2(i). Assume that up to time T, there exists a unique and almost surely well-defined solution to the equation*

$$Z_t^* = Z_0^* + \sum_{r\in\mathcal{R}^s} \xi_r^* Y_r \left(\int_0^t \lambda_r^*(Z_s^*)ds\right) + \sum_{r\in\mathcal{R}^f} \xi_r^* \int_0^t \lambda_r^*(Z_s^*)ds, \tag{4.27}$$

where the functions λ_r^* are the limit functions (4.26). Then, if $\widehat{Z}_0^N$ converges in probability to Z_0^*, the process $\widehat{Z}_{\cdot\wedge T}^N$ converges in probability to $Z_{\cdot\wedge T}^*$ with respect to the Skorohod distance.

Proof. Just note that, in our setting, $Z_\cdot^N$ is the process associated to a *single-scale system* satisfying the condition of Lemma 2.8 in Pfaffelhuber and Popovic [2013], and the result follows. □

Example 4.7. Consider again Example 4.1. In Example 4.4, we saw that the reduced SRN approximate the behaviour of (4.1) for N large, in the sense of Theorem 4.3 and Corollary 4.4. Here we present a weak limit for the process of the reduced reaction network, given by Lemma 4.6. It is easy to check that the probabilities π_{13}^N, π_{14}^N, π_{23}^N and π_{24}^N tend to 2/3, 1/3, 1/3 and 2/3, respectively, for $N \to \infty$. The weak limit is given by the deterministic system

$$\begin{cases} x_t(E) = x_0(E) \\ x_t(R) = x_0(R) + x_0(E)\displaystyle\int_0^t \Big(k_3 x_s(P_2) - (k_1+k_2)x_s(R)\Big)ds \\ x_t(P_1) = x_0(P_1) + x_0(E)\displaystyle\int_0^t \frac{2k_1+k_2}{3} x_s(R)ds \\ x_t(P_2) = x_0(P_2) + x_0(E)\displaystyle\int_0^t \left(\frac{k_1+2k_2}{3} x_s(R) - k_3 x_s(P_2)\right)ds, \end{cases}$$

where, according to the choice of α, the counts of the species E and the (scaled) concentrations of the species R, P_1, P_2 are considered.

Theorem 4.7. *Assume that the hypotheses of Lemma 4.6 are satisfied. Moreover, assume that both $\widehat{X}_0^N$ and $\widehat{Z}_0^N$ converge in probability to Z_0^*. Then, for any $\varepsilon > 0$,*

$$\sup_{t\in[0,T]} P\left(\left|\widehat{X}_t^N - Z_t^*\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0. \tag{4.28}$$

Moreover, for any continuous function $f\colon \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}} \to \mathbb{R}$ we have

$$P\left(\sup_{t\in[0,T]} \left|\int_0^t \Big(f(\widehat{X}_s^N) - f(Z_s^*)\Big)ds\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0. \tag{4.29}$$

Finally, if $\alpha > 0$ then

$$P\left(\sup_{t\in[0,T]}\left|\widehat{X}^N_t - Z^*_t\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0. \tag{4.30}$$

The latter gives weak convergence of $\widehat{X}^N_{\cdot\wedge T}$ to $Z^*_{\cdot\wedge T}$ in the Skorohod topology.

Proof. Since Z^* is almost surely unique and well defined, we have that for any $\nu > 0$, there exists a constant $\Psi_\nu > 0$ such that

$$P\left(\sup_{t\in[0,T]}|Z^*_t| > \Psi_\nu\right) < \nu.$$

Since the number of species is finite, due to (4.6a) there exists some constant $\Upsilon^*_\nu > 0$ such that for N large enough

$$P\left(\sup_{t\in[0,T]}\left|\widehat{X}^N_t\right| > \Upsilon^*_\nu\right) < \nu.$$

Let

$$\Psi^*_\nu = \max\left\{\Psi_\nu, \Upsilon^*_\nu\right\}.$$

Moreover, let $D(h)$ denote the disc of radius h in $\mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}_{\geq 0}$ centred in the origin, with respect to the euclidean norm. For any $r \in \mathcal{R}^0$, we define $\lambda^N_{b,r}(\cdot)$ such that

$$\lambda^N_{b,r}(x) = \begin{cases} \lambda^N_r(x) & \text{if } x \in D(\Psi^*_\nu) \\ (1-|x|-\Psi^*_\nu)\lambda^N_r\left(\frac{\Psi^*_\nu}{|x|}x\right) + (|x|-\Psi^*_\nu)N^{\beta_r} & \text{if } x \in D(\Psi^*_\nu+1)\setminus D(\Psi^*_\nu) \\ N^{\beta_r} & \text{otherwise.} \end{cases}$$

These functions are Lipschitz and define a new kinetics $\mathcal{K}^N_b$. Let $X^N_{b,\cdot}$, $Z^N_{b,\cdot}$ and $Z^*_{b,\cdot}$ be the corresponding processes, with

$$X^N_{b,0} = X^N_0 \mathbb{1}_{D(\Psi^*_\nu)}(\widehat{X}^N_0), \quad Z^N_{b,0} = Z^N_0 \mathbb{1}_{D(\Psi^*_\nu)}(\widehat{Z}^N_0) \quad \text{and} \quad Z^*_{b,0} = Z^*_0 \mathbb{1}_{D(\Psi^*_\nu)}(Z^*_0).$$

With this choice, we have

$$P\left(X^N_{b,0} = X^N_0\right) \geq 1-\nu, \quad P\left(Z^N_{b,0} = Z^N_0\right) \geq 1-\nu \quad \text{and} \quad P\left(Z^*_{b,0} = Z^*_0\right) \geq 1-\nu,$$

at least for N large enough (by hypothesis $\widehat{Z}^N_0$ converges in probability to Z^*_0). Therefore

$$P\left(\sup_{t\in[0,T]}\left|Z^*_{b,t}\right| > \Psi^*_\nu\right) \leq P\left(\sup_{t\in[0,T]}|Z^*_t| > \Psi^*_\nu\right) + \nu < 2\nu,$$
$$P\left(\sup_{t\in[0,T]}\left|\widehat{Z}^N_t\right| > \Psi^*_\nu\right) \leq P\left(\sup_{t\in[0,T]}\left|\widehat{Z}^N_{b,t}\right| > \Psi^*_\nu\right) + \nu.$$

The rates $\lambda^N_{b,r}(\cdot)$ satisfy the condition in Theorem 4.3 and

$$E\left[\left|\widehat{X}^N_{b,0} - \widehat{Z}^N_{b,0}\right|\right] \xrightarrow[N\to\infty]{} 0.$$

From Theorem 4.3, we have

$$\sup_{t\in[0,T]} P\left(\left|\widehat{X}^N_{b,t} - \widehat{Z}^N_{b,t}\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0,$$

and by Lemma 4.6,

$$P\left(\sup_{t\in[0,T]}\left|\widehat{Z}^N_t\right| > \Psi^*_\nu\right) \leq P\left(\sup_{t\in[0,T]}\left|\widehat{Z}^N_{b,t}\right| > \Psi^*_\nu\right) + \nu \xrightarrow[N\to\infty]{} P\left(\sup_{t\in[0,T]}\left|Z^*_{b,t}\right| > \Psi^*_\nu\right) + \nu < 3\nu.$$

Putting it all together, we have

$$\begin{aligned}
&\limsup_{N\to\infty}\sup_{t\in[0,T]} P\left(\left|\widehat{X}^N_t-\widehat{Z}^N_t\right|>\varepsilon\right)\\
&\qquad\le \limsup_{N\to\infty}\sup_{t\in[0,T]} P\left(\left|\widehat{X}^N_t-\widehat{Z}^N_t\right|>\varepsilon,\ \sup_{t\in[0,T]}\left(\left|\widehat{X}^N_t\right|\vee\left|\widehat{Z}^N_t\right|\right)>\Psi^*_\nu\right)+\\
&\qquad+\limsup_{N\to\infty}\sup_{t\in[0,T]} P\left(\left|\widehat{X}^N_t-\widehat{Z}^N_t\right|>\varepsilon,\ \sup_{t\in[0,T]}\left(\left|\widehat{X}^N_t\right|\vee\left|\widehat{Z}^N_t\right|\right)\le\Psi^*_\nu\right)\\
&\qquad\le \limsup_{N\to\infty} P\left(\sup_{t\in[0,T]}\left(\left|\widehat{X}^N_t\right|\vee\left|\widehat{Z}^N_t\right|\right)>\Psi^*_\nu\right)+\limsup_{N\to\infty}\sup_{t\in[0,T]} P\left(\left|\widehat{X}^N_{b,t}-\widehat{Z}^N_{b,t}\right|>\varepsilon\right)<4\nu.
\end{aligned}$$

Since $\nu>0$ is arbitrary, we have (4.23). Similarly,

$$\begin{aligned}
&\limsup_{N\to\infty} P\left(\sup_{t\in[0,T]}\left|\int_0^t\left(f(\widehat{X}^N_s)-f(\widehat{Z}^N_s)\right)ds\right|>\varepsilon\right)\\
&\qquad\le \limsup_{N\to\infty} P\left(\sup_{t\in[0,T]}\left|\int_0^t\left(f(\widehat{X}^N_s)-f(\widehat{Z}^N_s)\right)ds\right|>\varepsilon,\ \sup_{t\in[0,T]}\left(\left|\widehat{X}^N_t\right|\vee\left|\widehat{Z}^N_t\right|\right)>\Psi^*_\nu\right)+\\
&\qquad+\limsup_{N\to\infty} P\left(\sup_{t\in[0,T]}\left|\int_0^t\left(f(\widehat{X}^N_s)-f(\widehat{Z}^N_s)\right)ds\right|>\varepsilon,\ \sup_{t\in[0,T]}\left(\left|\widehat{X}^N_t\right|\vee\left|\widehat{Z}^N_t\right|\right)\le\Psi^*_\nu\right)\\
&\le \limsup_{N\to\infty} P\left(\sup_{t\in[0,T]}\left(\left|\widehat{X}^N_t\right|\vee\left|\widehat{Z}^N_t\right|\right)>\Psi^*_\nu\right)+\limsup_{N\to\infty} P\left(\sup_{t\in[0,T]}\left|\int_0^t\left(f(\widehat{X}^N_{b,s})-f(\widehat{Z}^N_{b,s})\right)ds\right|>\varepsilon\right)<4\nu,
\end{aligned}$$

which implies that (4.24) holds. Since $\widehat{Z}^N_\cdot$ converges in probability to $Z^*_\cdot$ in the Skorohod space, by a version of the continuous mapping theorem [Hoffmann-Jørgensen, 1994, Section 5.4] it follows that

$$P\left(\sup_{t\in[0,T]}\left|\int_0^t\left(f(\widehat{Z}^N_s)-f(Z^*_s)\right)ds\right|>\varepsilon\right)\xrightarrow[N\to\infty]{}0,$$

where we used that the Skorohod distance for continuous functions is equivalent to the uniform distance. Hence, (4.29) is a consequence of triangular inequality. By similar arguments and by Theorem 4.5, if $\alpha>0$ we have

$$P\left(\sup_{t\in[0,T]}\left|\widehat{X}^N_t-\widehat{Z}^N_t\right|>\varepsilon\right)\le 4\nu+P\left(\sup_{t\in[0,T]}\left|\widehat{X}^N_{b,t}-\widehat{Z}^N_{b,t}\right|>\varepsilon\right)\xrightarrow[N\to\infty]{}4\nu.$$

If $\alpha>0$ then $Z^*_\cdot$ is continuous, therefore by Lemma 4.6 we have

$$P\left(\sup_{t\in[0,T]}\left|\widehat{Z}^N_t-Z^*_t\right|>\varepsilon\right)\xrightarrow[N\to\infty]{}0.$$

The proof is then concluded by the arbitrariness of ν, and the triangular inequality. □

Remark 4.7. Convergence of the processes $\widehat{X}^N_{\cdot\wedge T}$ to the process $Z^*_{\cdot\wedge T}$ in occupation measure is implied by (4.29) [Kallenberg, 1974, Theorem 4.5].

Corollary 4.8. *Assume that the hypotheses of Lemma 4.6 are satisfied. Then, the difference between the processes $\widehat{X}^N_{\cdot\wedge T}$ and $Z^*_{\cdot\wedge T}$ converges in finite dimensional distribution to 0.*

Proof. The proof is identical to the proof of Corollary 4.4. Indeed, from Theorem 4.7, we have that (4.28) holds for any $\varepsilon>0$. Thus, for any finite set of time points $\{t_m\}_{m=0}^p\subseteq[0,T]$, we have that

$$\begin{aligned}
P\left(\max_{0\le m\le p}\left|\widehat{X}^N_{t_m}-Z^*_{t_m}\right|>\varepsilon\right)&=P\left(\bigcup_{m=0}^p\left\{\left|\widehat{X}^N_{t_m}-Z^*_{t_m}\right|>\varepsilon\right\}\right)\\
&\le\sum_{m=0}^p P\left(\left|\widehat{X}^N_{t_m}-Z^*_{t_m}\right|>\varepsilon\right)\xrightarrow[N\to\infty]{}0,
\end{aligned}$$

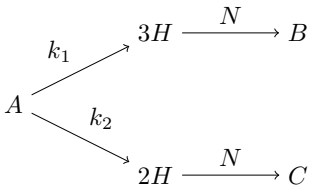
and the result follows. □

5 Discussion

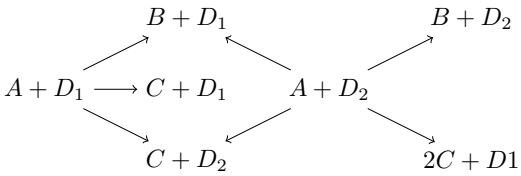
We close by presenting a collection of examples and remarks. A particular strength of our approach is that the reduced reaction network is easily found from the original reaction network and that the reaction rates of the reduced SRN can be found through a simple algebraic procedure. If the definition of intermediate species is relaxed, it might still be possible to find an approximating reduced SRN in concrete cases. However, a general technique does not seem to present itself easily.

We assume mass-action kinetics unless otherwise specified. If the stoichiometric coefficient of the intermediates were allowed to be different from one, or if different intermediate species were allowed to interact, our results would not be true in general:

Example 5.1 (Relaxing the definition of intermediates, I). Consider the SRN



with $\alpha = 0$. A single molecule of H could be trapped as the two reactions $3H \to B$ and $2H \to A$ compete against each other. Thus, there does not exist an approximation without intermediates as in Theorem 4.3 or Theorem 4.7. An approximation with no fast species, however, still exists. Since the dynamics of the system changes depending on whether a molecule of H is present or not, we might introduce two dummy variables D_1 and D_2 with $D_1 + D_2 = 1$, and $D_1 = 1$ if and only if no molecules of H are present. Let $\widehat{p}$ denote the projection onto the space of non-dummy variables. The finite dimensional distributions of $p(X_\cdot^N)$ are approximated by the finite dimensional distributions of $\widehat{p}(Z_\cdot^N)$, where $Z_\cdot^N$ is the process associated with



for a suitable choice of kinetics and with initial conditions $X_0(D_1) = 1$ and $X_0(D_2) = 0$. A general reduction technique that can deal with examples of this kind is subject of further investigation. Similar arguments can be made if intermediate species are interacting, for example, if $3H$ and $2H$ are replaced by $H_1 + H_2$ and H_1, respectively.

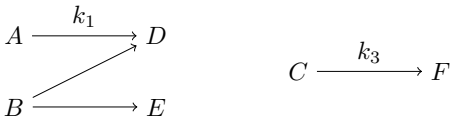
Example 5.2 (Relaxing the definition of intermediates, II). Consider the SRN below with $\alpha(C) = \alpha(F) = 1$ and $\alpha(A) = \alpha(B) = \alpha(D) = \alpha(E) = 0$:

$$A \xrightarrow{k_1} H_1 + H_2 \xrightarrow{N^7} D$$

$$B \xrightarrow{k_2} H_1 \xrightarrow{N} E$$

$$C \xrightarrow{k_3} H_2 \xrightarrow{N^2} F$$

Here, a reaction of type $C \to H_2$ can occur before a present molecule of H_1 is consumed, leading to the production of D from $H_1 + H_2 \to D$. It can be shown that the right limit is given by the rescaled process associated with



where

$$\lambda^N_{B\to E}(x) = k_2 x(B)\frac{N}{N + k_3 x(C)}, \qquad \lambda^N_{B\to D}(x) = k_2 x(B)\frac{k_3 x(C)}{N + k_3 x(C)}.$$

If we change the rate constant of $H_1 \to E$ to N^2 and let $\alpha(B) = \alpha(E) = 1$, a different reduced SRN is obtained in which a new complex appears:

$$A \xrightarrow{k_1} D \xleftarrow{\frac{k_2 k_3}{N^2}} B + C \qquad B \xrightarrow{k_2} E \qquad C \xrightarrow{k_3} F$$

It would be desirable to state Theorem 4.7 in terms of the stronger notion of convergence in probability in the Skorohod space, or at least in terms of the weak convergence in the Skorohod topology. This is done for $\alpha > 0$ (cf. Theorems 4.3 and 4.7), however it cannot be done in general as shown in the next example.

Example 5.3 (Weak convergence). Consider

$$A \xrightarrow{k} H \xrightarrow{N} B$$

with $\alpha(A) = \alpha(B) = 0$, and the limit process $\widehat{Z}^N_\cdot$ associated with the SRN

$$A \xrightarrow{k} B$$

Since the reduced SRN does not depend on N, we omit N in the notation. The possible states of $\widehat{Z}_\cdot$ satisfy the conservation law $\widehat{Z}_t(A) + \widehat{Z}_t(B) = M$ for some fixed M. In contrast, whenever the reaction $A \to H$ occurs in $X^N_\cdot$, for a short amount of time at least, $\widehat{X}^N_t(A) + \widehat{X}^N_t(B) \le M - 1$. The latter situation happens with positive probability, such that

$$E\left[\inf_{[0,T]}\left(\widehat{X}^N_t(A) + \widehat{X}^N_t(B)\right)\right] - E\left[\inf_{[0,T]}\left(\widehat{Z}_t(A) + \widehat{Z}_t(B)\right)\right] \xrightarrow[N\to\infty]{} c \neq 0.$$

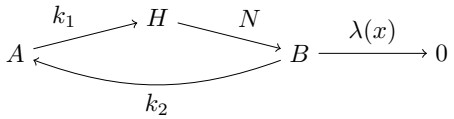
Hence, $\widehat{Z}_{\cdot\wedge T}$ does not provide a weak limit in the Skorohod topology for $\widehat{X}^N_{\cdot\wedge T}$. In fact, in this particular case the sequence of processes $\widehat{X}^N_{\cdot\wedge T}$ cannot have a weak limit in the Skorohod topology, since the sequence of the corresponding distributions P^N is not tight.

A natural question arising from the results of this paper is whether the reduced reaction network could be used to approximate the limit behaviour of the full model as $t \to \infty$. Specifically, we want to investigate whether for all Borel sets $A \subset \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$, it holds that

$$\lim_{t\to\infty} P\left(\widehat{X}^N_t \in A\right) - \lim_{t\to\infty} P\left(\widehat{Z}^N_t \in A\right) \xrightarrow[N\to\infty]{} 0, \tag{5.1}$$

under the hypothesis that the limits exist. The answer is negative, as it is shown with the next example.

Example 5.4 (Limit behaviour in the stochastic setting). Consider the following SRN.

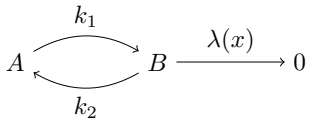


Let $\alpha(A) = \alpha(B) = 0$, assume that $X^N_0(A) + X^N_0(B) = M$ and $X^N_0(H) = 0$, and let

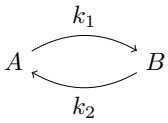
$$\lambda(x) = (M - x(A) - x(B))\mathbb{1}_{(0,\infty)}(x(B)).$$

The first occurrence of the reaction $B \to 0$ can only take place when H is present. Though it is unlikely for big N, there is still a positive probability that this happens, i.e. that $B \to 0$ occurs before the reaction $H \to B$ takes place. With probability one, all molecules of B will eventually be consumed, and the limit distribution of the above SRN is therefore concentrated on the state 0.

The SRN satisfies the assumptions of Theorem 4.3 and those of Theorem 4.7. The reduced reaction network is given by



where the initial conditions are the same as in the bigger model. Since in the reduced SRN $\lambda(Z_t^N) = 0$ whenever $Z_t^N(A) + Z_t^N(B) = M$, then the reaction $B \to 0$ never occurs. This implies that the reduced SRN is equivalent to



The limit distribution of the above SRN is concentrated on the set $\{x\colon x(A) + x(B) = M\}$. Therefore it is clear that (5.1) does not hold in this case. However, the limit distribution of the latter SRN approximates the *quasi-stationary distribution* of the original SRN when N tends to infinity, if we condition on the event that the reaction $B \to 0$ has not taken place; see for example Anderson et al. [2010, 2014] for a discussion on stationary and quasi-stationary distributions in reaction network theory.

6 Proof of Theorem 3.1 and Proposition 4.2

This section is devoted to prove Theorem 3.1 and Proposition 4.2. First, recall the transition rate matrix (3.2). Consider the continuous time Markov chain C^x with state space $\mathcal{U} \sqcup \mathcal{V} \sqcup \mathcal{W}$ (disjoint union) and transition rate matrix given by

$$Q(x) = \left[\begin{array}{c|c|c} Q^x_{U,U} & Q^x_{U,V} & 0 \\ \hline 0 & Q_{V,V} & Q_{V,W} \\ \hline 0 & 0 & 0 \end{array}\right],$$

where $Q^x_{U,V}$ is defined by

$$q^x_{i\ell} = \lambda_{i\ell}(x)$$

for $i \in U$, $\ell \in V$, and

$$Q^x_{U,U} = \operatorname{diag}(-Q^x_{U,V} e),$$

where e denotes a vector of suitable length with all entries 1. Given a matrix M, we denote by M_i its i-th row. Note that the matrix

$$\mathcal{L}^x_i = \left[\begin{array}{c|c} -Q_{V,V} & -Q_{V,W} e \\ \hline -(Q^x_{U,V})_i & (Q^x_{U,V})_i e \end{array}\right]$$

is the transposed Laplacian matrix of the graph $\mathcal{G}^x_i$ defined in (3.7) (row sums are zero). Let $D^x(\cdot)$ denote the discrete time Markov chain embedded in $C^x(\cdot)$ and let $\mathbb{P}(x)$ be the corresponding transition probability matrix of D^x. For any $i \in U$, let

$$P^x_i(\cdot) = P(\cdot | D^x(0) = y_i), \quad \text{and} \quad E^x_i[\cdot] = E[\cdot | D^x(0) = y_i].$$

Moreover, let

$$d_i^x = \sum_{\ell \in V_i} \lambda_{i\ell}(x) = (Q_{U,V}^x)_i e,$$
$$d_\ell = \sum_{\ell' \in V} k_{\ell\ell'} + \sum_{j \in W} k_{\ell j} = [Q_{V,V} \,|\, Q_{V,W}]_\ell \, e.$$

We have the following result:

Lemma 6.1. *For all $\ell \in V$, $i \in U$ and $x \in \mathbb{R}_{\geq 0}^{\mathcal{X}\setminus\mathcal{V}}$,*

$$\mu_{i\ell}(x) = \frac{d_i^x}{d_\ell} \sum_{n \geq 1} (\mathbb{P}(x)^n)_{i\ell} < \infty. \tag{6.1}$$

In particular, we have

$$\mu_{i\ell}(x) = \frac{d_i^x}{d_\ell} E_i^x \left[\#\textit{visits of } D^x(\cdot) \textit{ to } H_\ell \right].$$

Proof. We have

$$E_i^x \left[\#\text{visits of } D^x(\cdot) \text{ to } H_\ell \right] = E_i^x \left[\sum_{n \geq 1} \mathbb{1}_{\{H_\ell\}} D^x(n) \right] = \sum_{n \geq 1} P_i^x \left(D^x(n) = H_\ell \right) = \sum_{n \geq 1} (\mathbb{P}(x)^n)_{i\ell}.$$

Therefore, since every intermediate species is a transient state in D^x,

$$\sum_{n \geq 1} (\mathbb{P}(x)^n)_{i\ell} = E_i^x \left[\#\text{visits of } D^x(\cdot) \text{ to } H_\ell \right] < \infty.$$

Thus, we only need to prove (6.1). The matrix $\mathbb{P}(x)$ has the following block structure:

$$\mathbb{P}(x) = \left[\begin{array}{c|c|c} 0 & \mathbb{P}_{U,V}^x & 0 \\ \hline 0 & \mathbb{P}_{V,V} & \mathbb{P}_{V,U} \\ \hline 0 & 0 & I \end{array}\right].$$

Thus, we have

$$\mathbb{P}(x)^n = \left[\begin{array}{c|c|c} 0 & \mathbb{P}_{U,V}^x \mathbb{P}_{V,V}^{n-1} & * \\ \hline 0 & \mathbb{P}_{V,V}^n & * \\ \hline 0 & 0 & * \end{array}\right].$$

Since for any $\ell, \ell' \in V$,

$$\left(\sum_{n \geq 0} \mathbb{P}_{V,V}^n \right)_{\ell\ell'} = E \left[\#\text{visits of } D^x(\cdot) \text{ to } H_{\ell'} \,|D^x(1) = H_\ell \right] < \infty,$$

we have that $\sum_{n \geq 0} \mathbb{P}_{V,V}^n$ is well defined and

$$\sum_{n \geq 0} \mathbb{P}_{V,V}^n = (I - \mathbb{P}_{V,V})^{-1}.$$

Therefore

$$\sum_{n \geq 1} (\mathbb{P}(x)^n)_{i\ell} = \sum_{n \geq 0} (\mathbb{P}_{U,V}^x \mathbb{P}_{V,V}^n)_{i\ell} = \left(\mathbb{P}_{U,V}^x (I - \mathbb{P}_{V,V})^{-1} \right)_{i\ell}. \tag{6.2}$$

Assume $d_i^x \neq 0$. Consider the graph $\widetilde{\mathcal{G}}_i^x$ with the same nodes and edges as $\mathcal{G}_i^x$ and normalized labels

$$H_\ell \xrightarrow{\frac{k_{\ell\ell'}}{d_\ell}} H_{\ell'} \quad , \quad H_\ell \xrightarrow{\frac{\sum_{j\in W} k_{\ell j}}{d_\ell}} \star \quad , \quad \star \xrightarrow{\frac{\lambda_{i\ell}(x)}{d_i^x}} H_\ell.$$

The transpose of the Laplacian matrix of the graph $\widetilde{\mathcal{G}}_i^x$ is given by

$$\widetilde{\mathcal{L}}_i^x = \left[\begin{array}{c|c} I - \mathbb{P}_{V,V} & -\mathbb{P}_{V,W} e \\ \hline -(\mathbb{P}_{U,V}^x)_i & 1 \end{array}\right].$$

Given a matrix M, denote by $M_{(i,j)}$ the matrix obtained by M eliminating the i-th row and the j-th column. We have that

$$\begin{aligned}
\left(\mathbb{P}_{U,V}^x (I - \mathbb{P}_{V,V})^{-1}\right)_{i\ell} &= \sum_{\ell'\in V} (\mathbb{P}_{U,V}^x)_{i\ell'} (I - \mathbb{P}_{V,V})^{-1}_{\ell'\ell} = \sum_{\ell'\in V} (\mathbb{P}_{U,V}^x)_{i\ell'} \frac{(-1)^{\ell+\ell'} \det(I - \mathbb{P}_{V,V})_{(\ell,\ell')}}{\det(I - \mathbb{P}_{V,V})} \\
&= (-1)^{\ell+\#\mathcal{V}+1} \frac{\det(\widetilde{\mathcal{L}}_i^x)_{(\ell,\#\mathcal{V}+1)}}{\det(\widetilde{\mathcal{L}}_i^x)_{(\#\mathcal{V}+1,\#\mathcal{V}+1)}} = \frac{\det(\widetilde{\mathcal{L}}_i^x)_{(\ell,\ell)}}{\det(\widetilde{\mathcal{L}}_i^x)_{(\#\mathcal{V}+1,\#\mathcal{V}+1)}} \\
&= \frac{d_\ell}{d_i^x} \frac{\det(\mathcal{L}_i^x)_{(\ell,\ell)}}{\det(\mathcal{L}_i^x)_{(\#\mathcal{V}+1,\#\mathcal{V}+1)}} = \frac{d_\ell}{d_i^x} \frac{\sum_{\zeta\in\Theta_{i,x}(H_\ell)} w(\zeta)}{\sum_{\zeta\in\Theta_{i,x}(\star)} w(\zeta)} = \frac{d_\ell}{d_i^x} \mu_{i\ell}(x),
\end{aligned}$$

where the second equality follows from the co-factor expansion of the determinant, the third from the Laplace expansion and the fourth equality follows from the fact that the last column of the Laplacian matrix is equal to minus the sum of the other columns. The second-last equality follows from the Matrix-Tree theorem [Tutte, 1948]. Thus, from (6.2) it follows that (6.1) holds. If $d_i^x = 0$, then $\mu_{i\ell} = 0$ for all $\ell \in V$. Thus, (6.1) still holds and the proof is concluded. □

The proof of Theorem 3.1 follows from Lemma 6.1.

Proof of Theorem 3.1. We have to prove that for any fixed $x \in \mathbb{R}_{\geq 0}^{\mathcal{X}\setminus\mathcal{V}}$,

$$\sum_{\ell\in V} \pi_{\ell j} \lambda_{i\ell}(x) = \sum_{\ell\in V} k_{\ell j} \mu_{i\ell}(x). \tag{6.3}$$

Note that

$$\begin{aligned}
\sum_{\ell\in V} \pi_{\ell j} \lambda_{i\ell}(x) &= d_i^x \sum_{\ell\in V} P_i^x \left(\lim_{n\to\infty} D^x(n) = y_j \,|D^x(1) = H_\ell\right) P_i^x \left(D^x(1) = H_\ell\right) \\
&= d_i^x P_i^x \left(\lim_{n\to\infty} D^x(n) = y_j\right) = d_i^x \sum_{n\geq 1} \sum_{\ell\in V} P_i^x \left(D^x(n) = H_\ell \,,\, D^x(n+1) = y_j\right) \\
&= d_i^x \sum_{\ell\in V} \frac{k_{\ell j}}{d_\ell} \sum_{n\geq 1} (\mathbb{P}(x)^n)_{i\ell}.
\end{aligned}$$

Therefore, (6.3) follows from Lemma 6.1. □

To prove Proposition 4.2, we make the dependence on N explicit.

Proof of Proposition 4.2. From Lemma 6.1 we have that

$$\mu_{i\ell}^N(x) = \frac{d_i^{N,x}}{d_\ell^N} E_i^{N,x} \left[\#\text{visits of } D^{N,x}(\cdot) \text{ to } H_\ell\right]$$

for $x \in \mathbb{R}_{\geq 0}^{\mathcal{X}\setminus\mathcal{V}}$. Denote by T_ℓ^N the random variable distributed as the time until consumption of a molecule of H_ℓ. Its distribution is exponential with parameter d_ℓ^N. Note that the Markov chain $D^{N,x}(\cdot)$

is distributed as the discrete time Markov chain embedded in $C_1^N(\cdot)$, and $D^{N,x}(0)$ denotes the initial reactant setting off the chain $C_1^N(\cdot)$. For any $j \in W$, we have

$$d_i^{N,N^\alpha x} E\left[\tau_1^N \left| D^{N,N^\alpha x}(0) = y_i\right.\right] = d_i^{N,N^\alpha x} \sum_{\ell \in V} E_i^{N,N^\alpha x}\left[\#\text{visits of } D^{N,N^\alpha x}(\cdot) \text{ to } H_\ell\right] E\left[T_\ell^N\right]$$
$$= \sum_{\ell \in V} \mu_{i\ell}^N(N^\alpha x).$$

Furthermore,

$$d_i^{N,N^\alpha x} E\left[\tau_1^N \left| D^{N,N^\alpha x}(0) = y_i\right.\right] =$$
$$= \sum_{\ell \in V} E\left[\tau_1^N \left| D^{N,N^\alpha x}(1) = H_\ell\right.\right] P\left(D^{N,N^\alpha x}(1) = H_\ell \left| D^{N,N^\alpha x}(0) = y_i\right.\right) d_i^{N,N^\alpha x}$$
$$= \sum_{\ell \in V} E\left[\tau_1^N \left| D^{N,N^\alpha x}(1) = H_\ell\right.\right] \lambda_{i\ell}^N(N^\alpha x)$$
$$= \sum_{\ell \in V} E\left[N^{\beta_{i\ell}} \tau_1^N \left| D^{N,N^\alpha x}(1) = H_\ell\right.\right] N^{-\beta_{i\ell}} \lambda_{i\ell}^N(N^\alpha x).$$

In particular,

$$\sum_{\ell \in V} N^{-\alpha^*} \mu_{i\ell}^N(N^\alpha x) = \sum_{\ell \in V} N^{\beta_{i\ell} - \alpha^*} E\left[\tau_1^N \left| D^{N,N^\alpha x}(1) = H_\ell\right.\right] N^{-\beta_{i\ell}} \lambda_{i\ell}^N(N^\alpha x). \tag{6.4}$$

Therefore, (4.16) holds if and only if the right-hand side of (6.4) tends to zero as $N \to \infty$. By Assumption 2(iii) we have

$$N^{-\beta_{i\ell}} \lambda_{i\ell}^N(N^\alpha x) \xrightarrow[N\to\infty]{} \lambda_{i\ell}(x),$$

where $\lambda_{i\ell}$ is a non-null function. It follows that the right-hand side of (6.4) tends to zero as $N \to \infty$ if and only if, for any $i \in U, \ell \in V$, such that $y_i \to H_\ell \in \mathcal{R}$, and for any $j \in W$

$$N^{\beta_{i\ell} - \alpha^*} E\left[\tau_1^N \left| D^{N,N^\alpha x}(1) = H_\ell\right.\right] \xrightarrow[N\to\infty]{} 0.$$

The latter is equivalent to

$$N^{\beta_{i\ell} - \alpha^*} E\left[\tau_1^N \left| C_1^N(t_1^N) = H_\ell\right.\right] \xrightarrow[N\to\infty]{} 0. \tag{6.5}$$

By the definition of β_ℓ^* and a_ℓ, the latter implies

$$N^{\beta_\ell^* - a_\ell} E\left[\tau_1^N \left| C_1^N(t_1^N) = H_\ell\right.\right] \xrightarrow[N\to\infty]{} 0$$

for any $\ell \in \bigcup_{i\in U} V_i$, which is what we wanted to prove. If $\alpha_j^* = \alpha_{j'}^*$ for any $j, j' \in W$, then $a_\ell = \alpha^*$ for any $\ell \in V$. Therefore, (6.5) for any $j \in W$ is equivalent to (4.15). The proof is thus concluded. □

7 Proof of Theorems 4.3 and 4.5

In this section Theorems 4.3 and 4.5 are proven. To this aim, instead of assuming (4.10) and (4.11) in Assumption 2(v), we make use of the weaker conditions (4.12) and (4.13). Throughout this section, whenever t is written it is implicitly assume that $t \in [0, T]$. We also use the notation

$$\|x\|_\infty = \max_{S \in \mathcal{X}} |x(S)| \quad \text{for } x \in \mathbb{R}^{\mathcal{X}}.$$

By the equivalence of norms in $\mathbb{R}^{\mathcal{X}}$, we have that there exists $\theta > 0$, such that

$$|x| \leq \theta \|x\|_\infty \quad \forall x \in \mathbb{R}^{\mathcal{X}}.$$

Let $D_\infty(h)$ be the disc of radius h in $\mathbb{R}_{\geq 0}^{\mathcal{X}\setminus\mathcal{V}}$ centred in the origin, with respect to the measure $\|\cdot\|_\infty$, and let $D_\infty^C(r)$ be its complementary.
We start by stating a lemma.

Lemma 7.1. *Assume the assumptions of Theorem 4.3 hold. Then,*

$$\sup_{t\in[0,T]} N^{-\alpha_j^*} E\left[M_{i\ell j}^N(t) - \overline{M}_{i\ell j}^N(t)\right] \xrightarrow[N\to\infty]{} 0$$

Proof. Remember that $t_{i\ell j,n}^N$ is the time of the n-th jump of $M_{i\ell j}^N(t)$, and $\tau_{i\ell j,n}^N$ the life time of the corresponding chain of intermediates. Note that by (4.12) we have

$$0 \le \sup_{t\in[0,T]} N^{-\alpha_j^*} E\left[M_{i\ell j}^N(t) - \overline{M}_{i\ell j}^N(t)\right] \le \sup_{t\in[0,T]} N^{-\alpha_j^*} E\left[M_{i\ell j}^N(t)\right] \le \pi_{\ell j}^N N^{\beta_\ell^*-\alpha_j^*} B_{i\ell} t \le c_{\ell j} B_{i\ell} t.$$

This implies that the sequence $\sup_{t\in[0,T]} N^{-\alpha_j^*} E\left[M_{i\ell j}^N(t) - \overline{M}_{i\ell j}^N(t)\right]$ is contained in a compact set, and it follows that to prove the lemma it is sufficient to show that all the accumulation points of the sequence are 0. To this aim, fix an accumulation point l and consider a subsequence N_h such that

$$\sup_{t\in[0,T]} N_h^{-\alpha_j^*} E\left[M_{i\ell j}^{N_h}(t) - \overline{M}_{i\ell j}^{N_h}(t)\right] \xrightarrow[h\to\infty]{} l.$$

First, assume that

$$\liminf_{h\to\infty} \pi_{\ell j}^{N_h} N_h^{\beta_\ell^*-\alpha_j^*} = 0,$$

and let N_{h_m} a subsequence such that

$$\lim_{m\to\infty} \pi_{\ell j}^{N_{h_m}} N_{h_m}^{\beta_\ell^*-\alpha_j^*} = 0.$$

In this case,

$$0 \le l = \lim_{m\to\infty} \sup_{t\in[0,T]} N_{h_m}^{-\alpha_j^*} E\left[M_{i\ell j}^{N_{h_m}}(t) - \overline{M}_{i\ell j}^{N_{h_m}}(t)\right] \le \lim_{m\to\infty} \pi_{\ell j}^{N_{h_m}} N_{h_m}^{\beta_\ell^*-\alpha_j^*} B_{i\ell} t = 0,$$

which proves $l = 0$. Now, assume that

$$\liminf_{h\to\infty} \pi_{\ell j}^{N_h} N_h^{\beta_\ell^*-\alpha_j^*} = \delta > 0,$$

and fix $0 < \varepsilon < \delta t$. For convenience, denote

$$\sigma_{j\ell}^{\varepsilon h} = \frac{\varepsilon N_h^{\alpha_j^*-\beta_\ell^*}}{\pi_{\ell j}^{N_h}}.$$

We have

$$\begin{aligned}
E\left[M_{i\ell j}^{N_h}(t) - \overline{M}_{i\ell j}^{N_h}(t)\right] =& E\left[\sum_{n=1}^{M_{i\ell j}^{N_h}(t)} \mathbb{1}_{\left[t-t_{i\ell j,n}^{N_h},\infty\right)}(\tau_{i\ell j,n}^{N_h})\right] \\
\le& E\left[\sum_{n=1}^{M_{i\ell j}^{N_h}\left(t-\sigma_{j\ell}^{\varepsilon h}\right)} \mathbb{1}_{\left[\sigma_{j\ell}^{\varepsilon h},\infty\right)}(\tau_{i\ell j,n}^{N_h})\right] + E\left[\sum_{n=M_{i\ell j}^{N_h}\left(t-\sigma_{j\ell}^{\varepsilon h}\right)+1}^{M_{i\ell j}^{N_h}(t)} 1\right] \\
\le& E\left[\sum_{n=1}^{M_{i\ell j}^{N_h}(t)} \mathbb{1}_{\left[\sigma_{j\ell}^{\varepsilon h},\infty\right)}(\tau_{i\ell j,n}^{N_h})\right] + E\left[M_{i\ell j}^{N_h}(t) - M_{i\ell j}^{N_h}\left(t-\sigma_{j\ell}^{\varepsilon h}\right)\right].
\end{aligned}$$

Thus, using (4.17) and (4.21) we obtain

$$\begin{aligned}
E\left[M_{i\ell j}^{N_h}(t) - \overline{M}_{i\ell j}^{N_h}(t)\right] \le& E\left[\sum_{n=1}^{Y_{i\ell j}(tB_{i\ell}\pi_{\ell j}^{N_h} N_h^{\beta_{i\ell}})} \mathbb{1}_{\left[\sigma_{j\ell}^{\varepsilon h},\infty\right)}(\tau_{i\ell j,n}^N)\right] + \varepsilon N^{\alpha_j^*-\beta_\ell^*} N^{\beta_\ell^*} B_{i\ell} \\
\le& t p_{\ell j}^{\varepsilon}(N_h) N_h^{\beta_\ell^*} B_{i\ell} \pi_{\ell j}^{N_h} + \varepsilon N_h^{\alpha_j^*} B_{i\ell} \le B_{i\ell} N_h^{\alpha_j^*} (t\pi_{\ell j}^{N_h} N_h^{\beta_\ell^*-\alpha_j^*} p_{\ell j}^{\varepsilon}(N_h) + \varepsilon),
\end{aligned}$$

where $p^{\varepsilon}_{\ell j}(N)$ is as defined in (4.8). By (4.9) and the arbitrariness of $\varepsilon > 0$, the latter implies that

$$\sup_{t\in[0,T]} N_h^{-\alpha^*_j} E\left[M^{N_h}_{i\ell j}(t) - \overline{M}^{N_h}_{i\ell j}(t)\right] \xrightarrow[h\to\infty]{} 0,$$

which implies that $l = 0$ and concludes the proof. □

Proof of Theorem 4.3. Let the process $\widehat{W}^N_\cdot$ be defined as in (4.20) and, for any fixed t, let $\Delta^N_t = \left\|\widehat{X}^N_t - \widehat{W}^N_t\right\|_\infty$. Then, we have

$$\begin{aligned}
E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right] &\le E\left[\left|\widehat{W}^N_t - \widehat{Z}^N_t\right| + \left|\widehat{X}^N_t - \widehat{W}^N_t\right|\right] \\
&\le \sum_{r\in\mathcal{R}^1} \left|N^{-\alpha}\xi_r\right| E\left[\left|Y_r\left(\int_0^t \lambda^N_r(X^N_s)ds\right) - Y_r\left(\int_0^t \lambda^N_r(Z^N_s)ds\right)\right|\right] + \\
&\quad + \sum_{i\in U}\sum_{j\in W} \left|N^{-\alpha}(y_j - y_i)\right| \sum_{\ell\in V_i} E\left[\left|Y_{i\ell j}\left(\int_0^t \pi^N_{\ell j}\lambda^N_{i\ell}(X^N_s)ds\right) - Y_{i\ell j}\left(\int_0^t \pi^N_{\ell j}\lambda^N_{i\ell}(Z^N_s)ds\right)\right|\right] + \\
&\quad + E\left[\left|\widehat{X}^N_0 - \widehat{Z}^N_0\right|\right] + \theta E\left[\Delta^N_t\right] \\
&= \sum_{r\in\mathcal{R}^1} \left|N^{-\alpha+\beta_r}\xi_r\right| E\left[N^{-\beta_r}\left|Y_r\left(\int_0^t \lambda^N_r(X^N_s)ds\right) - Y_r\left(\int_0^t \lambda^N_r(Z^N_s)ds\right)\right|\right] + \\
&\quad + \sum_{i,j,\ell} \left|N^{-\alpha+\beta_{i\ell}}(y_j - y_i)\right| E\left[N^{-\beta_{i\ell}}\left|Y_{i\ell j}\left(\int_0^t \pi^N_{\ell j}\lambda^N_{i\ell}(X^N_s)ds\right) - Y_{i\ell j}\left(\int_0^t \pi^N_{\ell j}\lambda^N_{i\ell}(Z^N_s)ds\right)\right|\right] + \\
&\quad + E\left[\left|\widehat{X}^N_0 - \widehat{Z}^N_0\right|\right] + \theta E\left[\Delta^N_t\right].
\end{aligned}$$

For any reaction $r\colon y_i \to y_j \in \mathcal{R}^*$, let

$$\alpha^*_{ij} = \alpha^*_r = \min_{S\colon \xi_r(S)\neq 0} \alpha(S).$$

Then,

$$\begin{aligned}
E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right] &\le \sum_{r\in\mathcal{R}^1} N^{-\alpha^*_r+\beta_r}\left|\xi_r\right| E\left[\int_0^t N^{-\beta_r}\left|\lambda^N_r(N^\alpha \widehat{X}^N_s) - \lambda^N_r(N^\alpha \widehat{Z}^N_s)\right| ds\right] + \\
&\quad + \sum_{i,j,\ell} \pi^N_{\ell j} N^{-\alpha^*_{ij}+\beta_{i\ell}}\left|y_j - y_i\right| E\left[\int_0^t N^{-\beta_{i\ell}}\left|\lambda^N_{i\ell}(N^\alpha \widehat{X}^N_s) - \lambda^N_{i\ell}(N^\alpha \widehat{Z}^N_s)\right| ds\right] + \\
&\quad + E\left[\left|\widehat{X}^N_0 - \widehat{Z}^N_0\right|\right] + \theta E\left[\Delta^N_t\right].
\end{aligned}$$

To control the left side, we aim to substitute the functions $\lambda^N_r(\cdot)$ with their limits $\lambda_r(\cdot)$ (Assumption 2(iii)). To meet our goal, we first argue that the processes $\widehat{X}^N_\cdot$ and $\widehat{Z}^N_\cdot$ are bounded with high probability.

Let $S \in \mathcal{X}\setminus\mathcal{V}$. By substituting the rate functions $\lambda^N_r(\cdot)$ with their upper bounds $N^{\beta_r}B_r$ in (4.19), and using $t < T$, we obtain that $\widehat{Z}^N_t(S)$ is bounded from above by

$$\widehat{Z}^N_0(S) + \sum_{r\in\mathcal{R}^1_S} \left|\xi_r(S)\right| N^{-\alpha(S)} Y_r(N^{\beta_r}B_r T) + \sum_{i\in U}\sum_{j\in W} \left|y_j(S) - y_i(S)\right| \sum_{\ell\in V_i} N^{-\alpha(S)} Y_{i\ell j}\left(\pi^N_{\ell j} N^{\beta_{i\ell}} B_{i\ell} T\right),$$

where $\mathcal{R}^1_S$ is defined according to (2.2). Using the assumptions (4.12), (4.13) and the Law of Large Numbers for Poisson processes to control the above expression for $\alpha(S) > 0$, we obtain that, for any $\nu > 0$, there exists $\Upsilon'_\nu > 0$, such that

$$\limsup_{N\to\infty} P\left(\sup_{t\in[0,T]} \left\|\widehat{Z}^N_t\right\|_\infty > \Upsilon'_\nu\right) < \nu.$$

Let Υ_ν be as in (4.6a) and let $\Upsilon''_\nu = \Upsilon_\nu \vee \Upsilon'_\nu$. Then, if N is large enough,

$$P\left(\sup_{t\in[0,T]}\left(\left\|\widehat{X}^N_t\right\|_\infty \vee \left\|\widehat{Z}^N_t\right\|_\infty\right) > \Upsilon''_\nu\right) < 3\nu. \tag{7.1}$$

By Assumption 2(iii) we have that

$$N^{-\beta_r}\lambda^N_r(N^\alpha x) \xrightarrow[N\to\infty]{} \lambda_r(x) \quad \forall r \in \mathcal{R}^0$$

uniformly on compact sets. In particular, for any $\nu > 0$,

$$o_\nu(N) = \sup_{x\in D_\infty(\Upsilon''_\nu)} \left|N^{-\beta_r}\lambda^N_r(N^\alpha x) - \lambda_r(x)\right| \xrightarrow[N\to\infty]{} 0.$$

Note that for any $\nu > 0$ and $x \in \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}_{\geq 0}$,

$$\left|N^{-\beta_r}\lambda^N_r(N^\alpha x) - \lambda_r(x)\right| \leq o_\nu(N)\mathbb{1}_{D_\infty(\Upsilon''_\nu)}(x) + 2B_r\mathbb{1}_{D^C_\infty(\Upsilon''_\nu)}(x). \tag{7.2}$$

Using (7.1) and (7.2) we have

$$\begin{aligned}
E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right] \leq& \sum_{r\in\mathcal{R}^1} N^{-\alpha^*_r+\beta_r}|\xi_r|\left(E\left[\int_0^t\left|\lambda_r(\widehat{X}^N_s) - \lambda_r(\widehat{Z}^N_s)\right|ds\right] + 2o_\nu(N)t + 12B_r\nu t\right) + \\
&+ \sum_{i,j,\ell}\pi^N_{\ell j}N^{-\alpha^*_{ij}+\beta_{i\ell}}|y_j - y_i|\left(E\left[\int_0^t\left|\lambda_{i\ell}(\widehat{X}^N_s) - \lambda_{i\ell}(\widehat{Z}^N_s)\right|ds\right] + 2o_\nu(N)t + 12B_{i\ell}\nu t\right) + \\
&+ E\left[\left|\widehat{X}^N_0 - \widehat{Z}^N_0\right|\right] + \theta E\left[\Delta^N_t\right] \\
\leq& \Psi_1\int_0^t E\left[\left|\widehat{X}^N_s - \widehat{Z}^N_s\right|\right]ds + \Psi_2 o_\nu(N)t + \Psi_3\nu t + E\left[\left|\widehat{X}^N_0 - \widehat{Z}^N_0\right|\right] + \theta E\left[\Delta^N_t\right]
\end{aligned}$$

for some positive constants Ψ_1, Ψ_2, $\Psi_3 > 0$, independent of ν. In the last inequality we made use of (4.12) and (4.13), as well as the hypothesis that λ_r is Lipschitz for any $r \in \mathcal{R}^0$.

To prove (4.22), we only need to show that $\sup_{t\in[0,T]} E\left[\Delta^N_t\right] \to 0$ for $N \to \infty$. Indeed if this holds, then by the Gronwall inequality applied to the function $\sup_{t\in[0,T]} E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right]$ we have

$$\sup_{t\in[0,T]} E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right] \leq \left(\Psi_2 o_\nu(N)T + \Psi_3\nu T + E\left[\left|\widehat{X}^N_0 - \widehat{Z}^N_0\right|\right] + \theta\sup_{t\in[0,T]} E\left[\Delta^N_t\right]\right)e^{\Psi_1 T},$$

which for $N \to \infty$ tends to $\Psi_3\nu T e^{\Psi_1 T}$. By the arbitrariness of ν this leads to

$$\sup_{t\in[0,T]} E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right] \xrightarrow[N\to\infty]{} 0,$$

and we are done. To prove that $\sup_{t\in[0,T]} E\left[\Delta^N_t\right] \to 0$ for $N \to \infty$, it first follows from (4.18) and (4.20) that

$$\Delta^N_t = \left\|N^{-\alpha}\sum_{j\in W} y_j \sum_{i\in U}\sum_{\ell\in V_i}\left(M^N_{i\ell j}(t) - \overline{M}^N_{i\ell j}(t)\right)\right\|_\infty \leq \sum_{i,\ell,j}\|y_j\|_\infty N^{-\alpha^*_j}\left(M^N_{i\ell j}(t) - \overline{M}^N_{i\ell j}(t)\right). \tag{7.3}$$

Therefore, by Lemma 7.1 and (7.3), we have that $\sup_{t\in[0,T]} E\left[\Delta^N_t\right] \to 0$ for $N \to \infty$, which concludes the proof of the first part of the statement. Equation (4.23) is implied by (4.22) and the Markov inequality.

Finally, to prove (4.24), first consider a continuously differentiable function $g\colon \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}} \to \mathbb{R}$ with compact domain, and let c_g be the maximum of the absolute value of its derivative. We have

$$\begin{aligned}
\sup_{t\in[0,T]} E\left[\int_0^t\left|g(\widehat{X}^N_s) - g(\widehat{Z}^N_s)\right|ds\right] &= \int_0^T E\left[\left|g(\widehat{X}^N_s) - g(\widehat{Z}^N_s)\right|\right]ds \\
&\leq \int_0^T c_g E\left[\left|\widehat{X}^N_s - \widehat{Z}^N_s\right|\right]ds \\
&\leq T c_g \sup_{t\in[0,T]} E\left[\left|\widehat{X}^N_t - \widehat{Z}^N_t\right|\right]ds \xrightarrow[N\to\infty]{} 0.
\end{aligned}$$

Let $f\colon \mathbb{R}^{\mathcal{X}\setminus\mathcal{V}} \to \mathbb{R}$ be a continuous function with compact domain. By uniformly approximating f by continuously differentiable functions with compact domain, we have

$$\sup_{t\in[0,T]} E\left[\int_0^t \left|f(\widehat{X}_s^N) - f(\widehat{Z}_s^N)\right| ds\right] \xrightarrow[N\to\infty]{} 0.$$

By Markov inequality, it follows that for any $\varepsilon > 0$

$$\sup_{t\in[0,T]} P\left(\left|\int_0^t \left(f(\widehat{X}_s^N) - f(\widehat{Z}_s^N)\right) ds\right| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0. \tag{7.4}$$

Consider the occupation measures on $[0,T]\times\mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$ given by

$$\Gamma_1^N([t_1,t_2]\times A) = \int_{t_1}^{t_2} \mathbb{1}_A(\widehat{X}_s^N) ds \quad \text{and} \quad \Gamma_2^N([t_1,t_2]\times A) = \int_{t_1}^{t_2} \mathbb{1}_A(\widehat{Z}_s^N) ds$$

for any $0 \le t_1 < t_2 \le T$ and any Borel set A of $\mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}$. By (7.1) and Kurtz [1992, Lemma 1.3], we have that the sequences of random measures (Γ_1^N) and (Γ_2^N) are relatively compact with respect to the Prohorov metric. By the continuous mapping theorem [Hoffmann-Jørgensen, 1994, Section 5.4], this in turn implies that the sequence of continuous processes

$$\left(\int_0^\cdot f(\widehat{X}_s^N) ds, \int_0^\cdot f(\widehat{Z}_s^N) ds\right) = \left(\int_0^\cdot \int_{\mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}} f(x) d\Gamma_1^N(s,x), \int_0^\cdot \int_{\mathbb{R}^{\mathcal{X}\setminus\mathcal{V}}} f(x) d\Gamma_2^N(s,x)\right)$$

is relatively compact with respect to the weak convergence in the product space $D[0,T]\times D[0,T]$, where $D[0,T]$ denotes the usual Skorohod space. In this case it coincides with weak convergence in the uniform topology since the processes are continuous. The following is inspired by a proof of Donnelly and Kurtz [1996, Lemma A2.1]. Consider a weak limit $(\widehat{X}_\cdot, \widehat{Z}_\cdot)$, which will be a continuous process. By (7.4), we have $\widehat{X}_t = \widehat{Z}_t$ for any t, therefore $d_{Sk}(\widehat{X}_\cdot, \widehat{Z}_\cdot) = 0$, where d_{Sk} denotes the Skorohod distance. By the continuous mapping theorem, we have that for any subsequence converging to $(\widehat{X}_\cdot, \widehat{Z}_\cdot)$,

$$d_{Sk}\left(\int_0^\cdot f(\widehat{X}_s^{N_m}) ds, \int_0^\cdot f(\widehat{Z}_s^{N_m}) ds\right)$$

converges weakly to zero. In particular, this implies that for every $\varepsilon > 0$

$$P\left(\sup_{t\in[0,T]} \left|\int_0^t \left(f(\widehat{X}_s^{N_m}) - f(\widehat{Z}_s^{N_m})\right) ds\right| > \varepsilon\right) \xrightarrow[m\to\infty]{} 0.$$

Since the same holds for any convergent subsequence and by relative compactness, (4.24) follows for any continuous f with compact support. Indeed, if it were not the case we would have a subsequence such that for some constant $c > 0$

$$P\left(\sup_{t\in[0,T]} \left|\int_0^t \left(f(\widehat{X}_s^{N_m}) - f(\widehat{Z}_s^{N_m})\right) ds\right| > \varepsilon\right) > c.$$

However, the subsequence would not contain any convergent subsequence.

Now, let the support of f be not compact, and consider $\nu > 0$. There exists a continuous function f_ν with compact support such that $f_\nu(x) = f(x)$ if $\|x\|_\infty \le \Upsilon''_\nu$, where Υ''_ν is as in (7.1). Therefore, if N is large enough

$$\begin{aligned} &P\left(\sup_{t\in[0,T]} \left|\int_0^t \left(f(\widehat{X}_s^N) - f(\widehat{Z}_s^N)\right) ds\right| > \varepsilon\right) \\ &\qquad \le P\left(\sup_{t\in[0,T]} \left|\int_0^t \left(f_\nu(\widehat{X}_s^N) - f_\nu(\widehat{Z}_s^N)\right) ds\right| > \varepsilon\right) + 3\nu \xrightarrow[N\to\infty]{} 3\nu. \end{aligned}$$

The proof is concluded by the arbitrariness of ν. □

Proof of Theorem 4.5. Fix $\varepsilon > 0$. Let $\delta > 0$ be such

$$\left(\sum_{r \in \mathcal{R}_S^1} \|\xi_r\|_\infty B_r + \sum_{i \in U} \sum_{j \in W} \|y_j + y_i\|_\infty \sum_{\ell \in V_i} c_{\ell j} B_{i\ell} \right) \delta < \frac{\varepsilon}{3\theta}, \tag{7.5}$$

where $c_{\ell j}$ is as in (4.12). Now consider a sequence of real numbers $t_0 < t_1 < t_2 < \cdots < t_q$ such that $t_{m+1} - t_m < \delta$ for any $0 \leq m < q$, $t_0 = 0$ and $t_q = T$. For any $0 \leq m < q$ and any species $S \in \mathcal{X} \setminus \mathcal{V}$ we have

$$\begin{aligned} \sup_{t \in [t_m, t_{m+1}]} \left| \widehat{Z}_t^N(S) - \widehat{Z}_{t_m}^N(S) \right| &\leq N^{-\alpha(S)} \sum_{r \in \mathcal{R}_S^1} |\xi_r(S)| \left| Y_r(N^{\beta_r} B_r t_{m+1}) - Y_r(N^{\beta_r} B_r t_m) \right| + \\ &+ N^{-\alpha(S)} \sum_{i \in U} \sum_{j \in W} |y_j(S) - y_i(S)| \sum_{\ell \in V_i} \left| Y_{i\ell j} \left(\pi_{\ell j}^N N^{\beta_{i\ell}} B_{i\ell} t_{m+1} \right) - Y_{i\ell j} \left(\pi_{\ell j}^N N^{\beta_{i\ell}} B_{i\ell} t_m \right) \right|. \end{aligned}$$

The latter is distributed as

$$\begin{aligned} N^{-\alpha(S)} \sum_{r \in \mathcal{R}_S^1} |\xi_r(S)| \, Y_r \left(N^{\beta_r} B_r (t_{m+1} - t_m) \right) + \\ + N^{-\alpha(S)} \sum_{i \in U} \sum_{j \in W} |y_j(S) - y_i(S)| \sum_{\ell \in V_i} Y_{i\ell j} \left(\pi_{\ell j}^N N^{\beta_{i\ell}} B_{i\ell} (t_{m+1} - t_m) \right), \end{aligned}$$

which, due to $\alpha(S) > 0$, (4.12), (4.13), the Law of Large Numbers for Poisson processes and (7.5), is asymptotically bounded in probability by a constant strictly smaller than $\varepsilon/(3\theta)$. In particular,

$$P \left(\sup_{t \in [t_m, t_{m+1}]} \left| \widehat{Z}_t^N - \widehat{Z}_{t_m}^N \right| > \frac{\varepsilon}{3} \right) \xrightarrow[N \to \infty]{} 0.$$

Similarly, by (4.18)

$$\begin{aligned} &\sup_{t \in [t_m, t_{m+1}]} \left| \widehat{X}_t^N(S) - \widehat{X}_{t_m}^N(S) \right| \leq N^{-\alpha(S)} \sum_{r \in \mathcal{R}_S^1} |\xi_r(S)| \left| Y_r(N^{\beta_r} B_r t_{m+1}) - Y_r(N^{\beta_r} B_r t_m) \right| + \\ &+ N^{-\alpha(S)} \sum_{i \in U} \sum_{j \in W} \left(y_j(S) \sum_{\ell \in V_i} \left| \overline{M}_{i\ell j}^N(t_{m+1}) - \overline{M}_{i\ell j}^N(t_m) \right| + y_i \sum_{\ell \in V_i} \left| M_{i\ell j}^N(t_{m+1}) - M_{i\ell j}^N(t_m) \right| \right) \\ &\leq N^{-\alpha(S)} \sum_{r \in \mathcal{R}_S^1} |\xi_r(S)| \left| Y_r(N^{\beta_r} B_r t_{m+1}) - Y_r(N^{\beta_r} B_r t_m) \right| + \\ &+ N^{-\alpha(S)} \sum_{i \in U} \sum_{j \in W} (y_j(S) + y_i(S)) \sum_{\ell \in V_i} \left| Y_{i\ell j} \left(\pi_{\ell j}^N N^{\beta_{i\ell}} B_{i\ell} t_{m+1} \right) - Y_{i\ell j} \left(\pi_{\ell j}^N N^{\beta_{i\ell}} B_{i\ell} t_m \right) \right| + \\ &+ y_j(S) \sum_{\ell \in V_i} \left(\left| \overline{M}_{i\ell j}^N(t_{m+1}) - M_{i\ell j}^N(t_{m+1}) \right| + \left| \overline{M}_{i\ell j}^N(t_m) - M_{i\ell j}^N(t_m) \right| \right). \end{aligned}$$

Again, due to $\alpha(S) > 0$, (4.12), (4.13), the Law of Large Numbers for Poisson processes, (7.5) and Lemma 7.1, the latter is asymptotically bounded in probability by a constant strictly smaller than $\varepsilon/(3\theta)$. Specifically,

$$P \left(\sup_{t \in [t_m, t_{m+1}]} \left| \widehat{X}_t^N - \widehat{X}_{t_m}^N \right| > \frac{\varepsilon}{3} \right) \xrightarrow[N \to \infty]{} 0.$$

We have

$$\begin{aligned} &P \left(\sup_{t \in [0,T]} \left| \widehat{X}_t^N - \widehat{Z}_t^N \right| > \varepsilon \right) \\ &\qquad \leq P \left(\max_{0 \leq m < q} \left(\widehat{X}_{t_m}^N - \widehat{Z}_{t_m}^N, \sup_{t \in [t_m, t_{m+1}]} \left| \widehat{Z}_t^N - \widehat{Z}_{t_m}^N \right|, \sup_{t \in [t_m, t_{m+1}]} \left| \widehat{X}_t^N - \widehat{X}_{t_m}^N \right| \right) > \frac{\varepsilon}{3} \right). \end{aligned} \tag{7.6}$$

Hence, the proof is concluded by Corollary 4.4, which is direct consequence of Theorem 4.3. □

Acknowledgements

We thank Elisenda Feliu for reading and commenting on early versions of this manuscript and an anonymous reviewer for providing valuable suggestions that lead to improvement of the results.

References

Thomas G Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972.

Thomas G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.*, 6(3):223–240, 1977/78. ISSN 0304-4149.

Karen Ball, Thomas G. Kurtz, Lea Popovic, and Greg Rempala. Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.*, 16(4):1925–1961, 2006. ISSN 1050-5164. doi: 10.1214/105051606000000420. URL http://dx.doi.org/10.1214/105051606000000420.

Hye-Won Kang and Thomas G. Kurtz. Separation of time-scales and model reduction for stochastic reaction networks. *Ann. Appl. Probab.*, 23(2):529–583, 2013. ISSN 1050-5164. doi: 10.1214/12-AAP841. URL http://dx.doi.org/10.1214/12-AAP841.

Peter Pfaffelhuber and Lea Popovic. Scaling limits of spatial chemical reaction networks. *arXiv preprint arXiv:1302.0774*, 2013.

A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, London, third edition, 2004.

Hanna M Härdin, Antonios Zagaris, Klaas Krab, and Hans V Westerhoff. Simplified yet highly accurate enzyme kinetics for cases of low substrate concentrations. *FEBS journal*, 276(19):5491–5506, 2009.

Philipp Thomas, Arthur V Straube, and Ramon Grima. The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. *BMC systems biology*, 6(1), 2012.

Christopher V Rao and Adam P Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: application to the gillespie algorithm. *The Journal of chemical physics*, 118(11):4999–5010, 2003.

Elisenda Feliu and Carsten Wiuf. Simplifying biochemical models with intermediate species. *Journal of The Royal Society Interface*, 10(87), 2013a.

Elisenda Feliu and Carsten Wiuf. Variable elimination in chemical reaction networks with mass-action kinetics. *SIAM J. Appl. Math.*, 72(4):959–981, 2012. ISSN 0036-1399. doi: 10.1137/110847305. URL http://dx.doi.org/10.1137/110847305.

Elisenda Feliu and Carsten Wiuf. Variable elimination in post-translational modification reaction networks with mass-action kinetics. *J. Math. Biol.*, 66(1-2):281–310, 2013b. ISSN 0303-6812. doi: 10.1007/s00285-012-0510-4. URL http://dx.doi.org/10.1007/s00285-012-0510-4.

P. Érdi and J. Tóth. *Mathematical models of chemical reactions*. Nonlinear Science: Theory and Applications. Princeton University Press, Princeton, NJ, 1989. ISBN 0-691-08532-3. Theory and applications of deterministic and stochastic models.

David F Anderson and Thomas G Kurtz. Continuous time markov chain models for chemical reaction networks. In *Design and analysis of biomolecular circuits*, pages 3–42. Springer, 2011.

F. Horn and R. Jackson. General mass action kinetics. *Arch. Rational Mech. Anal.*, 47:81–116, 1972. ISSN 0003-9527.

J. Hoffmann-Jørgensen. *Probability with a view toward statistics. Vol. I*. Chapman & Hall Probability Series. Chapman & Hall, New York, 1994. ISBN 0-412-05221-0. doi: 10.1007/978-1-4899-3019-4. URL http://dx.doi.org/10.1007/978-1-4899-3019-4.

Olav Kallenberg. *Lectures on random measures*. Citeseer, 1974.

David F. Anderson, Gheorghe Craciun, and Thomas G. Kurtz. Product-form stationary distributions for deficiency zero chemical reaction networks. *Bull. Math. Biol.*, 72(8):1947–1970, 2010. ISSN 0092-8240. doi: 10.1007/s11538-010-9517-4. URL http://dx.doi.org/10.1007/s11538-010-9517-4.

David F Anderson, Germán A Enciso, and Matthew D Johnston. Stochastic analysis of biochemical reaction networks with absolute concentration robustness. *Journal of The Royal Society Interface*, 11 (93), 2014.

W. T. Tutte. The dissection of equilateral triangles into equilateral triangles. volume 44, pages 463–482, 1948.

Thomas G. Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied stochastic analysis (New Brunswick, NJ, 1991)*, volume 177 of *Lecture Notes in Control and Inform. Sci.*, pages 186–209. Springer, Berlin, 1992. doi: 10.1007/BFb0007058. URL http://dx.doi.org/10.1007/BFb0007058.

Peter Donnelly and Thomas G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.*, 24(2):698–742, 1996. ISSN 0091-1798. doi: 10.1214/aop/1039639359. URL http://dx.doi.org/10.1214/aop/1039639359.

IA

Elimination of Intermediate Species in Deterministic Reaction Networks

A non-submitted result that connects Paper I to the deterministic setting is shown in this appendix. Specifically, in Paper I a reduced model is introduced in the setting of stochastic reaction systems by means of elimination of intermediate species. Various kinds of convergences are then studied. In this appendix, we show that a reduced model can be similarly introduced in the setting of deterministic reaction systems, and we study the convergence of the solutions of the full and reduced systems up to a fixed time T.

In this appendix we consider all the concentrations of the different species as properly rescaled counts, such that all the concentrations are of order $O(1)$. Hence, the multiscale framework of Paper I is abandoned. In particular, here we allow only the consumption rates of intermediate species to tend to infinity. In order to relate the results presented here with those of Paper I, we consider the latter in the particular case of $\alpha = 0$.

The notation of Paper I is mantained. However, here the deterministic reaction systems are taken into account, therefore a *kinetics* $\mathcal{K}$ will be a set of functions indexed by $\mathcal{R}$ of the form

$$\begin{array}{rccc} \lambda_r\colon & \mathbb{R}_{\geq 0}^{\mathcal{X}} & \to & \mathbb{R}_{\geq 0} \\ & x & \mapsto & \lambda_r(x). \end{array}$$

We allow reaction rates to be constantly 0, in which case the corresponding reaction could be removed from the network. Moreover, the entries of the vector $X_t \in \mathbb{R}^{\mathcal{X}}$ are in this case the concentrations of the various species at time t, and the evolution of X_t is modelled by the system of ordinary differential equations

$$\frac{dX_t}{dt} = \sum_{r\in\mathcal{R}} \xi_r \lambda_r\left(X_t\right), \tag{IA.1}$$

subject to the initial condition $X_0 = x_0$. We repeat here the definition of deterministic mass-action kinetics, which is given by

$$\lambda_{y_i\to y_j}(x) = k_{ij} \prod_{S\in y_i} x(S)^{y_i(S)} \mathbb{1}_{\{x(S)>0\}},$$

where the constants k_{ij} are non-negative real numbers called *rate constants*. We denote this by $y_i \xrightarrow{k_{ij}} y_j$.

Intermediate Species in the deterministic setting

We use the definition of intermediates given in Paper I, specifically we refer to Definition 2.1. Roughly speaking, the presence of intermediate species "slows down" any reaction path that proceeds through the formation of intermediates. This is intuitive, since the production and degradation of a sequence of intermediate species delays the synthesis of the final product. In a model without intermediates, this would happen instantaneously. Let $(\mathcal{X}, \mathcal{C}, \mathcal{R})$ a reaction network with a set of intermediate species $\mathcal{V} \subset \mathcal{X}$, and consider a kinetics $\mathcal{K}$. To investigate the effects of the presence of intermediate species in detail, we first make the following assumption, also assumed in Paper I:

Assumption IA.1 (Reaction rates and intermediates)**.** *The consumption of the intermediates is governed by mass-action kinetics, namely for any $\ell, \ell' \in V$ and $j \in W$,*

$$\lambda_{\ell j}(x) = k_{\ell j} x(H_\ell) \quad \textit{and} \quad \lambda_{\ell\ell'}(x) = k_{\ell\ell'} x(H_\ell),$$

for some non-negative constants $k_{\ell j}$, $k_{\ell\ell'}$. Further, we assume that all other reaction rates do not depend on H_ℓ.

As in Paper I, consider the labelled directed graph $\mathcal{G}_i^x$ with node set $\mathcal{V} \cup \{\star\}$ and labelled edge set given by:

$$\begin{aligned}
&\bullet\ H_\ell \xrightarrow{k_{\ell\ell'}} H_{\ell'} && \text{if } k_{\ell\ell'} \neq 0 \text{ and } \ell \neq \ell' \\
&\bullet\ H_\ell \xrightarrow{\sum_{j\in W} k_{\ell j}} \star && \text{if } \sum_{j\in W} k_{\ell j} \neq 0 \\
&\bullet\ \ \star \xrightarrow{\lambda_{i\ell}(x)} H_\ell && \text{if } \lambda_{i\ell}(x) \neq 0
\end{aligned} \tag{IA.2}$$

If we order the nodes of the graph such that $\star$ is the last one and we maintain the notation of Section 6 of Paper I, the Laplacian matrix of (IA.2) has the form

$$L_i^x = (\mathcal{L}_i^x)^T = \left[\begin{array}{c|c} -M & -(Q^x_{U,V})_i^T \\ \hline e^T M & (Q^x_{U,V})_i e \end{array}\right], \tag{IA.3}$$

where, for convenience, we introduced

$$M = Q^T_{V,V}.$$

Specifically, for any $\ell, \ell' \in V$,

$$M_{\ell\ell'} = \begin{cases} k_{\ell'\ell} & \text{if } \ell \neq \ell' \\ -\sum_{h\in V\cup W} k_{\ell h} & \text{if } \ell = \ell' \end{cases}, \qquad (Q^x_{U,V})_i(\ell) = \lambda_{i\ell}(x).$$

Let X_t be as in (IA.1), denote by $\widehat{X}_t$ the projection of X_t onto the non-intermediates species space, and let $\mathring{X}_t$ the projection onto the intermediate species space. Moreover, define the vector Λ_t of length $\#\mathcal{V}$, with entries

$$\Lambda_t^N(\ell) = \sum_{i\in U} \lambda_{i\ell}^N\left(\widehat{X}_t^N\right).$$

With this definition, the vector of intermediate species concentrations $\mathring{X}_t$ is solution to

$$\frac{d}{dt}\mathring{X}_t = M\mathring{X}_t + \Lambda_t,$$

which implies that the general solution for the intermediate species concentrations is given by

$$\mathring{X}_t = \exp(Mt)\,\mathring{X}_0 + \int_0^t \exp(M(t-s))\,\Lambda_s ds. \tag{IA.4}$$

The non-intermediate species evolve according to

$$\begin{aligned}\frac{d}{dt}\widehat{X}_t &= \sum_{\substack{\ell\in V\\ j\in W}} y_j k_{\ell j} X_t(H_\ell) + \sum_{\substack{y\in\mathcal{C}\setminus\mathcal{V}\\ y'\in\mathcal{C}}} p(\xi_{y\to y'})\,\lambda_{y\to y'}(\widehat{X}_t)\\ &= \sum_{j\in W} y_j \kappa_j \mathring{X}_t + \sum_{\substack{y\in\mathcal{C}\setminus\mathcal{V}\\ y'\in\mathcal{C}}} p(\xi_{y\to y'})\,\lambda_{y\to y'}(\widehat{X}_t).\end{aligned} \tag{IA.5}$$

where p is the projection onto the non-intermediate species space and κ_j are row vectors of length $\#\mathcal{V}$ defined by $\kappa_j(\ell) = k_{\ell j}$ for any $j \in W$. Under the assumption $\mathring{X}_0 = 0$, we have from (IA.4),

$$\frac{d}{dt}\widehat{X}_t = \sum_{j\in W} y_j \kappa_j \int_0^t \exp(M(t-s))\,\Lambda_s ds + \sum_{\substack{y\in\mathcal{C}\setminus\mathcal{V}\\ y'\in\mathcal{C}}} p(\xi_{y\to y'})\,\lambda_{y\to y'}(\widehat{X}_t). \tag{IA.6}$$

The above is a system of delayed differential equations with a distributed delay, in the sense of Kuang (1993). In particular, (IA.6) does not depend explicitly on the intermediate species counts.

The Reduced Reaction System

Consider a reaction network $(\mathcal{X}, \mathcal{C}, \mathcal{R})$ with a set of intermediate species $\mathcal{V} \subset \mathcal{X}$. We define the reduced reaction network exactly as in Paper I, i.e. as

$$(\mathcal{X}\setminus\mathcal{V}, \mathcal{C}\setminus\mathcal{V}, \mathcal{R}^*), \tag{IA.7}$$

where $\mathcal{R}^*$ consists of all the reactions $r : y_i \to y_j$ involving no intermediates, such that r was already an element of $\mathcal{R}$, or the complex y_j is obtainable from y_i through a chain of intermediates. When the original reaction network is equipped with a kinetics $\mathcal{K}$, we introduce a kinetics $\mathcal{K}^*$ for (IA.7) following the definition appearing in Feliu & Wiuf (2013). To introduce $\mathcal{K}^*$, we first need some other definitions. These are already given in Paper I, but we repeat them here for convenience.

Let the labelled directed graph $\mathcal{G}_i^x$ be as in (IA.2). Let $\Theta_i^x(\cdot)$ be the set of labelled spanning trees of $\mathcal{G}_i^x$ rooted at the argument, and let $w(\cdot)$ be the product of the edge labels of the tree in the argument. Then, define $\mu_{i\ell}(x)$ as

$$\mu_{i\ell}(x) = \frac{\sum_{\sigma\in\Theta_i^x(H_\ell)} w(\sigma)}{\sum_{\sigma\in\Theta_i^x(\star)} w(\sigma)}. \tag{IA.8}$$

If $\lambda_{i\ell}(x) = 0$ for some $x \in \mathbb{N}^{\mathcal{X}}$, then there might be no spanning tree rooted at a given intermediate species. In that case, $\mu_{i\ell}(x)$ is 0. The denominator is always strictly positive as any intermediate is eventually turned into a non-intermediate. The kinetics $\mathcal{K}^*$ is therefore defined by the rate functions

$$\lambda^*_{ij}(x) = \lambda_{ij}(x) + \sum_{\ell \in V} k_{\ell j} \mu_{i\ell}(x) \quad \text{for } y_i \to y_j \in \mathcal{R}^*. \tag{IA.9}$$

The definition of the reduced system coincides with the one given in Feliu & Wiuf (2013) and is equivalent to that of Paper I (see Paper I, Theorem 3.1). It is worth noting that the aim of Feliu & Wiuf (2013) is studying the number and type of steady states in the original model in comparison with those of the reduced one, while here we investigate the approximation of the dynamics of the original system up to a finite time T, by means of the reduced one.

The convergence result

Consider a reaction network $(\mathcal{X}, \mathcal{C}, \mathcal{R})$ with a set of intermediate species $\mathcal{V} \subset \mathcal{X}$. Let $\mathcal{K}^N$ a sequence of kinetics, indexed by $N \in \mathbb{N}$. Let X^N_t denote the solution of the deterministic reaction system associated with the kinetics $\mathcal{K}^N$, and assume Assumption IA.1 holds. Moreover, denote by $\widehat{X}^N_t$ the projection of X^N_t onto the non-intermediate species, that is $p(X^N_t)$, and let Z^N_t denote the solution of the reduced model, with $\widehat{X}^N_0 = Z^N_0$. In order to approximate $\widehat{X}^N_t$ by Z^N_t, we make two key hypotheses: the first states that for a fixed finite time $T > 0$, there exists a finite constant $\Upsilon > 0$, such that

$$\forall\, N \quad \sup_{[0,T]} \left(\left| \widehat{X}^N_t \right| + \left| Z^N_t \right| \right) < \Upsilon. \tag{IA.10}$$

Uniform boundedness of $\widehat{X}^N_t$ corresponds to equation (4.6a) of Assumption 2(ii) in Paper I. Sufficient conditions for (IA.10) are given in the following Proposition:

Proposition IA.1. *Assume that Assumption IA.1 holds and that the kinetics $\mathcal{K}^N$ are Lipschitz. If $\mathring{X}^N_0 = 0$ and there exists a finite constant $\Upsilon > 0$ such that $\left|\widehat{X}^N_0\right|, \left|Z^N_0\right| \leq \Upsilon$ for any N, then* (IA.10) *holds.*

The proof is given in Section "Proofs".

The second hypothesis relates to Assumption 2(iv): the variable τ^N_1 is a phase-type distribution with subgenerator matrix $Q^N_{V,V}$, therefore equation (4.14) for $\alpha = 0$, that is

$$P\left(\tau^N_1 > \varepsilon \middle| C^N_1(t^N_1) = H_\ell\right) \xrightarrow[N\to\infty]{} 0 \quad \forall \ell \in \bigcup_{i \in U} V_i,$$

is equivalent to

$$(e_\ell)^T \exp\left(\varepsilon Q^N_{V,V}\right) e \xrightarrow[N\to\infty]{} 0 \quad \forall \ell \in \bigcup_{i \in U} V_i.$$

Here we assume something slightly stronger, that is

$$e^T \exp\left(\varepsilon Q^N_{V,V}\right) e \xrightarrow[N\to\infty]{} 0. \tag{IA.11}$$

We are now ready to enunciate the following convergence result, which states uniform convergence of the solution of the reduced model to that of the full model, up to a fixed time T:

Theorem IA.2. *Let $\{\mathcal{X}, \mathcal{C}, \mathcal{R}\}$ be a reaction network and let $\mathcal{K}^N$ be a family of locally Lipschitz kinetics. Assume that Assumption IA.1 holds. Furthermore, assume that $\widehat{X}_0^N = Z_0^N$ for any N and that the initial concentrations of intermediate species are 0. Then, if* (IA.10) *and* (IA.11) *hold, we have that*

$$\sup_{t\in[0,T]} \left|\widehat{X}_t^N - Z_t^N\right| \xrightarrow[N\to\infty]{} 0.$$

The proof of the result is given in Section "Proofs".

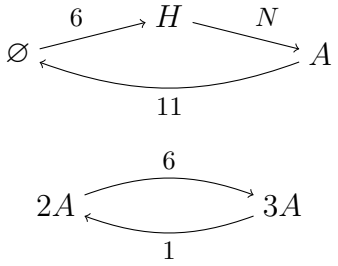
Limit behaviour

A natural question arising from Theorem IA.2 is whether the reduced reaction network could be used to approximate the limit behaviour of the full model as $t \to \infty$. Specifically, we inquire whether it holds that

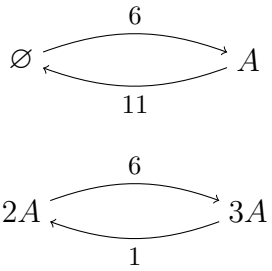
$$\lim_{t\to\infty} \widehat{X}_t^N - \lim_{t\to\infty} Z_t^N \xrightarrow[N\to\infty]{} 0, \tag{IA.12}$$

under the hypothesis that the above limits exist. The answer is that (IA.12) may not hold. Consider for example the case when Z_0^N is an unstable equilibrium point for the reduced model. Then $\lim_{t\to\infty} Z_t^N = Z_0^N$, while in the full reaction network with intermediates, a small perturbation given by the presence of intermediate species may push $\widehat{X}_t^N$ away from the repulsive point $X_0^N = Z_0^N$.

Consider the following deterministic mass-action system:



The assumptions of Theorem IA.2 are fulfilled and the reduced network is given by



The ODE governing the dynamics of the reduced network, which does not depend on N, is given by

$$\begin{aligned}\frac{dZ_t(A)}{dt} &= -Z_t(A)^3 + 6Z_t(A)^2 - 11Z_t(A) + 6 \\ &= -\left(Z_t(A) - 1\right)\left(Z_t(A) - 2\right)\left(Z_t(A) - 3\right) \doteqdot f\left(Z_t(A)\right).\end{aligned}$$

(IA.90)

Note that 2 is an unstable equilibrium point of the above dynamical system. We will show that if we assume $\widehat{X}_0^N(A) = Z_0(A) = 2$ and $X_0^N(H) = 0$, then (IA.12) does not hold.
The ODE system governing the dynamics of the full model is

$$\begin{cases} \dfrac{dX_t^N(H)}{dt} = 6 - NX_t^N(H) \\ \dfrac{dX_t^N(A)}{dt} = -X_t^N(A)^3 + 6X_t^N(A)^2 - 11X_t^N(A) + NX_t^N(H) \end{cases}.$$

This means that $X_t^N(H) = 6(1 - e^{-Nt})/N$ and

$$\frac{d\widehat{X}_t^N(A)}{dt} = -\widehat{X}_t^N(A)^3 + 6\widehat{X}_t^N(A)^2 - 11\widehat{X}_t^N(A) + 6\left(1 - e^{-Nt}\right) \doteqdot g_t^N\left(\widehat{X}_t^N(A)\right)$$

Since for any $N \geq 1$ and any $t > 0$ we have that $g_t^N(x) < f(x)$, and since $f(x) < 0$ for any $x \in (1, 2)$, then

$$\lim_{t\to\infty} \widehat{X}_t^N(A) \leq 1.$$

It is possible to prove a more precise result, namely that $\lim_{t\to\infty} \widehat{X}_t^N(A) = 1$, and it is worth noting that 1 is a stable steady state of the reduced model. Since

$$\lim_{t\to\infty} Z_t(A) = Z_0(A) = 2,$$

we have that (IA.12) does not hold.

Proofs

This section is devoted to the proofs of Proposition IA.1 and Theorem IA.2. Before proving the statements, we need some preliminary results. The first lemma is concerned with some properties of the matrix $\exp\left(Q_{V,V}^N t\right)$.

Lemma IA.3. *For any $t > 0$, any entry of the matrix $\exp\left(Q_{V,V}^N t\right)$ is non-negative. Moreover for $0 \leq s \leq t$, we have that*

$$\exp\left(Q_{V,V}^N s\right) e \geq \exp\left(Q_{V,V}^N t\right) e.$$

In particular, for $s = 0$ and any $t > 0$, $\exp\left(Q_{V,V}^N t\right) e \leq e$.

Proof. If we put $\Lambda_t^N \equiv 0$, then from (IA.4) we have

$$\mathring{X}_t^N = \exp\left(M^N t\right) \mathring{X}_0^N.$$

This implies that each row of $\exp\left(Q_{V,V}^N t\right)$ represents the concentrations of the intermediate species at time t given the initial condition $\mathring{X}_0^N = e_\ell$. In turn this implies that the row entries of $\exp\left(Q_{V,V}^N t\right)$ must be non-negative for any $t > 0$. Furthermore, the condition $\Lambda_s^N \equiv 0$ implies that the intermediates are not produced, thus the sum of their concentrations can only decrease. Indeed, the stoichiometric coefficients of the intermediate species are one, hence the net flow among intermediates is 0, while they can degrade to produce a non-intermediate complex. These considerations prove the second part of the lemma. □

Here the proof of Proposition IA.1is given.

Proof of Proposition IA.1. The Lipschitz condition implies that that there exists a constant Γ_1 such that for any $y \in \mathcal{C} \setminus \mathcal{V}$ and any $y' \in \mathcal{C}$

$$\lambda_{y \to y'}(x) \leq \Gamma_1 |x| .$$

Since by Theorem 3.1 in Paper I the rates for the reduced network given by (IA.9) can be written as

$$\lambda^*_{ij}(x) = \lambda_{ij}(x) + \sum_{\ell \in V_i} \pi_{\ell j} \lambda_{i\ell}(x),$$

where $\pi_{\ell j}$ are probabilities, it follows that there exists a positive constant Γ_2, such that, for $Z^N_t \neq 0$,

$$\frac{d\left|Z^N_t\right|}{dt} \leq \left|\frac{dZ^N_t}{dt}\right| \leq \Gamma_2 \left|Z^N_t\right| .$$

This implies

$$\left|Z^N_t\right| \leq \left|Z^N_0\right| e^{\Gamma_2 t}. \tag{IA.13}$$

Moreover, for any $N \in \mathbb{N}$ and any $\ell \in V$

$$\Lambda^N_t(\ell) \leq \#\mathcal{U} \cdot \Gamma_1 \left|\widehat{X}^N_t\right| .$$

By (IA.4) we have

$$\int_0^t \kappa^N_j \mathring{X}^N_v dv = \kappa^N_j \int_0^t \int_0^v \exp\left(M^N(v-s)\right) \Lambda^N_s ds dv.$$

Due to Lemma IA.3, all matrix and vector entries on the right-hand side are non-negative, thus we can use first the Fubini Theorem and then (IA.16) to obtain

$$\begin{aligned}
\int_0^t \kappa^N_j \mathring{X}^N_v dv &= \kappa^N_j \int_0^t \int_s^t \exp\left(M^N(v-s)\right) dv\, \Lambda^N_s ds \\
&\leq -e^T M^N \int_0^t \int_s^t \exp\left(M^N(v-s)\right) dv\, \Lambda^N_s ds \\
&= -e^T \int_0^t \Big(\exp\left(M^N(t-s)\right) - I \Big) \Lambda^N_s ds,
\end{aligned}$$

and by Lemma IA.3,

$$\leq e^T \int_0^t \Lambda^N_s ds \leq \#\mathcal{U} \cdot \#\mathcal{V} \cdot \Gamma_1 \int_0^t \left|\widehat{X}^N_s\right| ds.$$

Thus, by integrating (IA.5) we have that there exist some positive finite constants $\Gamma_3, \Gamma_4, \Gamma_5$, such that, for $\widehat{X}^N_t \neq 0$,

$$\left|\widehat{X}^N_t\right| \leq \Gamma_3 \sum_{j \in W} \int_0^t \kappa^N_j \mathring{X}^N_v dv + \Gamma_4 \int_0^t \left|\widehat{X}^N_v\right| dv \leq \Gamma_5 \int_0^t \left|\widehat{X}^N_v\right| dv,$$

which implies

$$\left|\widehat{X}^N_t\right| \leq \left|\widehat{X}^N_0\right| e^{\Gamma_5 t}. \tag{IA.14}$$

The conclusion of the proof now follows from (IA.13) and (IA.14), and from the hypothesis that there exists a finite constant $\Upsilon > 0$ bounding the initial conditions. □

Now we focus on the proof of Theorem IA.2. Without loss of generality we can assume that the kinetics $\mathcal{K}^N$ are Lipschitz: indeed, they are locally Lipschitz by hypothesis and their arguments are uniformly bounded by (IA.10). Using (IA.5), it is not hard to show that

$$\frac{d}{dt}\left(\widehat{X}_t^N - Z_t^N\right) = \sum_{j\in W} y_j \kappa_j^N \mathring{X}_t^N - \sum_{\substack{i\in U\\ j\in W}} y_j \sum_{\ell\in V} k_{\ell j}^N \mu_{i\ell}^N\left(Z_t^N\right) + \sum_{\substack{y\in\mathcal{C}\setminus\mathcal{V}\\ y'\in\mathcal{C}}} p\left(\xi_{y\to y'}\right)\left(\lambda_{y\to y'}\left(\widehat{X}_t^N\right) - \lambda_{y\to y'}\left(Z_t^N\right)\right).$$

In particular, using Cauchy-Schwartz and the Lipschitz condition, we have that for $\left|\widehat{X}_t^N - Z_t^N\right| \neq 0$,

$$\begin{aligned}\frac{d}{dt}\left|\widehat{X}_t^N - Z_t^N\right| &\leq \left|\frac{d}{dt}\left(\widehat{X}_t^N - Z_t^N\right)\right| \\ &\leq \left|\sum_{j\in W} y_j \kappa_j^N \mathring{X}_t^N - \sum_{\substack{i\in U\\ j\in W}} y_j \sum_{\ell\in V} k_{\ell j}^N \mu_{i\ell}^N\left(Z_t^N\right)\right| + \Gamma_0\left|\widehat{X}_t^N - Z_t^N\right| \qquad \text{(IA.15)}\end{aligned}$$

for some positive constant Γ_0.

We will make use of the following lemma.

Lemma IA.4. *Assume that Assumption IA.1 and (IA.10) hold, and that $\mathring{X}_0^N = 0$ for any $N \in \mathbb{N}$. Then there exist two positive constants Υ' and ρ, such that for any $N \in \mathbb{N}$ and for any $t \in [0,T]$,*

$$\left|\widehat{X}_t^N - Z_t^N\right| < \Upsilon'\left(e^{\rho t} - 1\right).$$

Moreover, there exists a positive constant Υ'', such that for any $s,t \in [0,T]$,

$$\left|\Lambda_t^N - \Lambda_s^N\right| \leq \Upsilon''|t-s|.$$

Proof. By hypothesis, the functions $\lambda_{i\ell}$ are continuous and their arguments are uniformly bounded since (IA.10) holds. Thus, there exists a positive constant B, such that for any $N \in \mathbb{N}$ and $\ell \in V$, we have

$$\sup_{t\in[0,T]}\left|\Lambda_\ell^N(t)\right| \leq B.$$

Moreover, we will show that for any $j \in W$, $\kappa_j^N \mathring{X}_t^N$ is uniformly bounded in $[0,T]$. For any $j \in W$,

$$-Q_{V,V}^N e = \left(\sum_{j\in W} k_{\ell j}^N\right)_{\ell\in V} \geq \left(\kappa_j^N\right)^T. \qquad \text{(IA.16)}$$

Lemma IA.3 gives that for any $t > 0$, the matrix $\exp\left(Q_{V,V}^N t\right)$ is non-negative. Using that $\exp\left(M^N t\right)$

and $\mathring{X}_t^N$ have non-negative entries, and using (IA.4) with $\mathring{X}_0^N = 0$, we obtain

$$\begin{aligned}
\kappa_j^N \mathring{X}_t^N &\leq -e^T M^N \mathring{X}_t^N \\
&= -e^T M^N \int_0^t \exp\left(M^N(t-s)\right) \Lambda_s^N ds \\
&\leq -e^T M^N \int_0^t \exp\left(M^N(t-s)\right) e\, ds \\
&= B e^T M^N \left(M^N\right)^{-1} \left(I - \exp\left(M^N t\right)\right) e \\
&= B\left(\#\mathcal{V} - e^T \exp\left(M^N t\right) e\right) \leq \#\mathcal{V} \cdot B,
\end{aligned}$$

where we have used Lemma IA.3 and the definition of M^N in the last inequality. Thus, by (IA.15) we have that, for $\left|\widehat{X}_t^N - Z_t^N\right| \neq 0$,

$$\frac{d}{dt}\left|\widehat{X}_t^N - Z_t^N\right| \leq \Gamma_1 \#\mathcal{V} \cdot B + \Gamma_2 B + \Gamma_0 \left|\widehat{X}_t^N - Z_t^N\right|$$

for some positive constants Γ_1, Γ_2. It follows that $\left|\widehat{X}_t^N - Z_t^N\right|$ is bounded by the solution to the differential equation corresponding to the right-hand side of the above inequality, with initial condition equal to 0, that is,

$$\left|\widehat{X}_t^N - Z_t^N\right| \leq \frac{B\left(\#\mathcal{V} \cdot \Gamma_1 + \Gamma_2\right)}{\Gamma_0}\left(e^{\Gamma_0 t} - 1\right).$$

This concludes the first part of the proof.

For any $t \in [0, T]$,

$$\left|\frac{d}{dt}\widehat{X}_t^N\right| \leq \Gamma_1 \sum_{j \in W} \kappa_j^N \mathring{X}_t^N + \Gamma_0 \left|\widehat{X}_t^N\right| \leq \Gamma_1 \#\mathcal{W} \cdot \#\mathcal{V} \cdot B + \Gamma_0 \Upsilon \doteq \overline{B}.$$

Thus, if the kinetics are Lipschitz, which we can assume without loss of generality, then for some positive constant Γ_4,

$$\left|\Lambda_t^N - \Lambda_s^N\right| \leq \Gamma_4 \left|\widehat{X}_t^N - \widehat{X}_s^N\right| \leq \Gamma_4 \overline{B}\left|t - s\right|,$$

which concludes the proof. □

To proceed we need another lemma. A similar result appears in Feliu & Wiuf (2013), and we give the proof here for completeness.

Lemma IA.5. *Let $\mu_{i\ell}^N(x)$ be defined as in (IA.8). We have that the l-th entry*

$$-\left(\left(M^N\right)^{-1} \Lambda_t^N\right)_\ell = \sum_{i \in U} \mu_{i\ell}^N\left(\widehat{X}_t^N\right).$$

Proof. The result does not depend on N, thus N is ignored in the notation. As discussed in the proof of Theorem 3.1 in Paper I, from the Matrix Tree Theorem we have that

$$\mu_{i\ell}\left(\widehat{X}_t\right) = \frac{\sum_{\sigma \in \theta_{i,\widehat{X}_t}(H_\ell)} w(\sigma)}{\sum_{\sigma \in \theta_{i,\widehat{X}_t}(\star)} w(\sigma)} = \frac{\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)_{(\ell,\ell)}}{\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)_{(\#\mathcal{V}+1,\#\mathcal{V}+1)}},$$

(IA.94)

where $\mathcal{L}_i^x$ is the transpose of the Laplacian matrix of the graph $\mathcal{G}_i^x$. Since the last column of $\mathcal{L}_i^x$ is the opposite of the sum of the other columns, we have

$$\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)_{(\ell,\ell)} = (-1)^{\ell+\#\mathcal{V}+1}\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)_{(\ell,\#\mathcal{V}+1)} = (-1)^{\ell+\#\mathcal{V}+1}\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)^T_{(\#\mathcal{V}+1,\ell)} \\ = -\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)^T_{\left(\#\mathcal{V}+1,\widehat{\ell}\right)},$$

where $(\mathcal{L}_i^x)^T_{\left(\#\mathcal{V}+1,\widehat{\ell}\right)}$ denotes the transpose of the matrix $\mathcal{L}_i^x$ with the $(\#\mathcal{V}+1)$-th row and column eliminated and the ℓ-th column replaced by the $(\#\mathcal{V}+1)$-th column. The last equality follows from changing the order of the columns in the transpose matrix. Moreover,

$$\mathcal{L}_i^x = \left[\begin{array}{c|c} -Q_{V,V} & -Q_{V,W}e \\ \hline -\left(Q^x_{U,V}\right)_i & \left(Q^x_{U,V}\right)_i e \end{array}\right].$$

Therefore, recalling that by definition $M = Q^T_{V,V}$, by Cramer's Rule we have

$$\mu_{i\ell}\left(\widehat{X}_t\right) = \frac{-\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)^T_{\left(\#\mathcal{V}+1,\widehat{\ell}\right)}}{\det\left(\mathcal{L}_i^{\widehat{X}_t}\right)^T_{(\#\mathcal{V}+1,\#\mathcal{V}+1)}} = -\left(M^{-1}\left(Q^{\widehat{X}_t}_{U,V}\right)_i\right)_\ell.$$

It follows that

$$\sum_{i\in U}\mu_{i\ell}\left(\widehat{X}_t\right) = -\left(M^{-1}\sum_{i\in U}\left(Q^{\widehat{X}_t}_{U,V}\right)_i\right)_\ell = -\left(M^{-1}\Lambda^N_t\right)_\ell,$$

which concludes the proof. □

We are now ready to prove Theorem IA.2.

Proof of Theorem IA.2. Let $\Upsilon', \Upsilon'', \rho > 0$ be as in Lemma IA.4. Fix $\varepsilon > 0$ and let $\delta_\varepsilon > 0$, such that

$$\sup_{t\in[0,\delta_\varepsilon]}\left|\widehat{X}^N_t - Z^N_t\right| < \Upsilon'\left(e^{\rho\delta_\varepsilon}-1\right) < \varepsilon. \tag{IA.17}$$

By (IA.15), it holds that, for any $t \in [0,T]$,

$$\begin{aligned}\frac{d}{dt}\left|\widehat{X}^N_t - Z^N_t\right| &\le \left|\sum_{\substack{\ell\in V\\ j\in W}} y_j k^N_{\ell j} X^N_t(H_\ell) - \sum_{\substack{i\in U\\ j\in W}} y_j \sum_{\ell\in V} k^N_{\ell j}\mu^N_{i\ell}\left(Z^N_t\right)\right| + \Gamma_0\left|\widehat{X}^N_t - Z^N_t\right| \\ &\le \left|\sum_{\substack{\ell\in V\\ j\in W}} y_j k^N_{\ell j} X^N_t(H_\ell) - \sum_{\substack{i\in U\\ j\in W}} y_j \sum_{\ell\in V} k^N_{\ell j}\mu^N_{i\ell}\left(\widehat{X}^N_t\right)\right| + \Gamma'_0\left|\widehat{X}^N_t - Z^N_t\right| \\ &= \left|\sum_{\substack{\ell\in V\\ j\in W}} y_j k^N_{\ell j} X^N_t(H_\ell) + \sum_{\substack{\ell\in V\\ j\in W}} y_j k^N_{\ell j}\left(\left(M^N\right)^{-1}\Lambda^N_t\right)_\ell\right| + \Gamma'_0\left|\widehat{X}^N_t - Z^N_t\right|,\end{aligned}$$

To obtain the second inequality we add and subtract $\sum_{\substack{i\in U\\ j\in W}} y_j \sum_{\ell\in V} k^N_{\ell j}\mu^N_{i\ell}\left(\widehat{X}^N_t\right)$ in the norm of the first term and use the Lipschitz condition again. Finally, the last equality comes from Lemma IA.5. For the sake of simplicity, we let

$$\beta^N(t) = \left| \sum_{\substack{\ell\in V\\ j\in W}} y_j k^N_{\ell j} X^N_t(H_\ell) + \sum_{\substack{\ell\in V\\ j\in W}} y_j k^N_{\ell j} \left(\left(M^N\right)^{-1} \Lambda^N_t \right)_\ell \right|.$$

From the above inequality, we have that for any $t \in [\delta_\varepsilon, T]$

$$\frac{d}{dt}\left|\widehat{X}^N_t - Z^N_t\right| \leq \sup_{s\in[\delta_\varepsilon,T]} \beta^N(s) + \Gamma'_0 \left|\widehat{X}^N_t - Z^N_t\right|.$$

Thus, if we let $\varphi\colon [\delta_\varepsilon, T] \to \mathbb{R}$ be the solution to the ODE system

$$\begin{cases} \dfrac{d\varphi(t)}{dt} = \sup\limits_{s\in[\delta_\varepsilon,T]} \beta^N(s) + \Gamma'_0\varphi(t) \\ \varphi(\delta_\varepsilon) = \left|\widehat{X}^N_{\delta_\varepsilon} - Z^N_{\delta_\varepsilon}\right|, \end{cases}$$

then, for any $t \in [\delta_\varepsilon, T]$, we have that $\left|\widehat{X}^N_t - Z^N_t\right| \leq \varphi(t)$. We solve the ODE above and obtain that for any $t \in [\delta_\varepsilon, T]$,

$$\left|\widehat{X}^N_t - Z^N_t\right| \leq e^{\Gamma'_0(t-\delta_\varepsilon)} \left(\left|\widehat{X}^N(\delta_\varepsilon) - Z^N(\delta_\varepsilon)\right| + \frac{\sup_{s\in[\delta_\varepsilon,T]} \beta^N(s)}{\Gamma'_0} \left(1 - e^{-\Gamma'_0(t-\delta_\varepsilon)}\right) \right).$$

In particular, from (IA.17) we have

$$\sup_{t\in[0,T]} \left|\widehat{X}^N_t - Z^N_t\right| \leq e^{\Gamma'_0 T} \left(\varepsilon + \frac{\sup_{s\in[\delta_\varepsilon,T]} \beta^N(s)}{\Gamma'_0} \right). \tag{IA.18}$$

If we can show that $\sup_{s\in[\delta_\varepsilon,T]} \beta^N(s)$ goes to zero with $N \to \infty$, then the theorem is proved. By definition of $\beta^N(t)$, this is equivalent to showing that for any $j \in W$,

$$\sup_{t\in[\delta_\varepsilon,T]} \kappa^N_j \left(\mathring{X}^N_t + \left(M^N\right)^{-1} \Lambda^N_t \right) \xrightarrow[N\to\infty]{} 0.$$

Using Lemma IA.4 we have

$$\begin{aligned} \mathring{X}^N_t &= \int_0^t \exp\left(M^N(t-s)\right) \Lambda^N_s ds \\ &= \int_0^t \exp\left(M^N(t-s)\right) \Lambda^N_t ds + \int_0^t \exp\left(M^N(t-s)\right) \left(\Lambda^N_s - \Lambda^N_t\right) ds \\ &\leq \left(M^N\right)^{-1} \left(\exp\left(M^N t\right) - I\right) \Lambda^N_t + \int_0^t \exp\left(M^N(t-s)\right) e\Upsilon''(t-s) ds. \end{aligned} \tag{IA.19}$$

Then, using the non-negativity of $\exp\left(M^N(t-s)\right)$, for any $\delta' \in (0,t)$, we have

$$\begin{aligned}
&\int_0^t \exp\left(M^N(t-s)\right) e\Upsilon''(t-s)ds \\
&\qquad \leq \int_0^{t-\delta'} \exp\left(M^N(t-s)\right) e\Upsilon'' T ds + \int_{t-\delta'}^{t} \exp\left(M^N(t-s)\right) e\Upsilon''\delta' ds \\
&\qquad \leq -\left(M^N\right)^{-1}\Big\{\left[\exp\left(M^N\delta'\right) - \exp\left(M^N t\right)\right] e\Upsilon'' T + \left[I - \exp\left(M^N\delta'\right)\right] e\Upsilon''\delta'\Big\}.
\end{aligned}$$

From (IA.16) it follows that, for any $j \in W$,

$$\kappa_j^N \leq -e^T M^N.$$

Therefore, using Lemma IA.3 and the definition of M^N for the second inequality , we have for any $\delta' \in (0, \delta_\varepsilon)$,

$$\begin{aligned}
&\sup_{t\in[\delta_\varepsilon,T]} \kappa_j^N \int_0^t \exp\left(M^N(t-s)\right) e\Upsilon''(t-s)ds \leq \\
&\qquad \leq \sup_{t\in[\delta_\varepsilon,T]} \Big(e^T\left[\exp\left(M^N\delta'\right) - \exp\left(M^N t\right)\right] e\Upsilon'' T + e^T\left[I - \exp\left(M^N\delta'\right)\right] e\Upsilon''\delta'\Big) \\
&\qquad \leq e^T\left[\exp\left(M^N\delta'\right) - \exp\left(M^N\delta_\varepsilon\right)\right] e\Upsilon'' T + e^T\left[I - \exp\left(M^N\delta'\right)\right] e\Upsilon''\delta'.
\end{aligned}$$

By (IA.11), when N tends to infinity the latter tends to

$$e^T e\Upsilon''\delta' = \#\mathcal{V}\cdot\Upsilon''\delta'.$$

Since $\delta' \in (0,\delta_\varepsilon)$ is arbitrary, it follows that

$$\sup_{t\in[\delta_\varepsilon,T]} \kappa_j^N \int_0^t \exp\left(M^N(t-s)\right) e\Upsilon''(t-s)ds \xrightarrow[N\to\infty]{} 0. \tag{IA.20}$$

Combining (IA.19) and (IA.20) and by making use of (IA.16), we have that for any $j \in W$,

$$\begin{aligned}
&\sup_{t\in[\delta_\varepsilon,T]} \kappa_j^N \left(\mathring{X}_t^N + \left(M^N\right)^{-1}\Lambda_t^N\right) \leq \\
&\qquad \leq \sup_{t\in[\delta_\varepsilon,T]} \kappa_j^N \left[\left(M^N\right)^{-1}\exp\left(M^N t\right)\Lambda_t^N + \int_0^t \exp\left(M^N(t-s)\right) e\Upsilon''(t-s)ds\right] \\
&\qquad \leq \sup_{t\in[\delta_\varepsilon,T]} \exp\left(M^N t\right)\Lambda_t^N + \sup_{t\in[\delta_\varepsilon,T]} \kappa_j^N \int_0^t \exp\left(M^N(t-s)\right) e\Upsilon''(t-s)ds \\
&\qquad \leq \sup_{t\in[\delta_\varepsilon,T]} \exp\left(M^N t\right) eB + \sup_{t\in[\delta_\varepsilon,T]} \kappa_j^N \int_0^t \exp\left(M^N(t-s)\right) e\Upsilon''(t-s)ds \xrightarrow[N\to\infty]{} 0.
\end{aligned}$$

Thus, we have proved that $\sup_{s\in[\delta_\varepsilon,T]}\beta^N(s)$ goes to zero with $N\to\infty$. From (IA.18) it follows that

$$\lim_{N\to\infty}\sup_{t\in[0,T]}\left|\widehat{X}_t^N - Z_t^N\right| \leq e^{\Gamma_0' T}\varepsilon,$$

and since $\varepsilon > 0$ is arbitrary, the proof is concluded. □

References

Feliu, E, & Wiuf, C. 2013. Simplifying biochemical models with intermediate species. *Journal of the royal society interface*, **10**(87).

Kuang, Y. 1993. *Delay differential equations: with applications in population dynamics*. Academic Press.

II

Product-form Poisson-like Distributions and Complex Balanced Reaction Systems

Daniele Cappelletti
Department of Mathematical Sciences
University of Copenhagen

Carsten Wiuf
Department of Mathematical Sciences
University of Copenhagen

Publication details

Submitted (2015).

Product-form Poisson-like distributions and complex balanced reaction systems

Daniele Cappelletti* Carsten Wiuf*

September 24, 2015

Abstract

Stochastic reaction networks are dynamical models of biochemical reaction systems and form a particular class of continuous-time Markov chains on $\mathbb{N}^n$. Here we provide a fundamental characterisation that connects structural properties of a network to its dynamical features. Specifically, we define the notion of 'stochastically complex balanced systems' in terms of the network's stationary distribution and provide a characterisation of stochastically complex balanced systems, parallel to that established in the 70-80ies for deterministic reaction networks. Additionally, we establish that a network is stochastically complex balanced if and only if an associated deterministic network is complex balanced (in the deterministic sense), thereby proving a strong link between the theory of stochastic and deterministic networks. Further, we prove a stochastic version of the 'deficiency zero theorem' and show that any (not only complex balanced) deficiency zero reaction network has a product-form Poisson-like stationary distribution on all irreducible components. Finally, we provide sufficient conditions for when a product-form Poisson-like distribution on a single (or all) component(s) implies the network is complex balanced, and explore the possibility to characterise complex balanced systems in terms of product-form Poisson-like stationary distributions.

1 Introduction

Improved experimental techniques have made it possible to measure molecular fluctuations at a small scale, creating a need for stochastic description of molecular data [23, 12]. Traditionally, biochemical reaction networks are modelled as deterministic systems of ordinary differential equations (ODEs), but these models assume the individual species are in high concentrations and do not allow for stochastic fluctuation. An alternative is stochastic models based on continuous-time Markov chains [14, 4, 12]. As an example of a stochastic reaction system, consider

$$A + B \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} 2C, \tag{1.1}$$

where κ_1, κ_2 are positive reaction constants. The network consists of three chemical species A, B and C and two reactions. Each occurrence of a reaction modifies the species counts, for example, when the reaction $A + B \to 2C$ takes places, the amount of A and B molecules are each decreased by one, while two molecules of C are created. The species counts are modelled as a continuous-time Markov chain, where the transitions are single occurrences of reactions with transition rates

$$\lambda_1(x) = \kappa_1 x_A x_B, \quad \lambda_2(x) = \kappa_2 x_C(x_C - 1),$$

and $x = (x_A, x_B, x_C)$ are the species counts [4]. When modelled deterministically, the concentrations (rather than the counts) of the species change according to an ODE system.

In a classical paper [18], Kurtz explored the relationship between deterministic and stochastic reaction systems, using a scaling argument – large volume limit – to link the dynamical behaviour of the two types of systems to each other. Other, mainly recent work, also points to close connections between the two types of systems [22, 2, 3, 1, 5, 16]. In this paper we explore this relationship further.

*Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark. The authors are supported by the Danish Research Councils.

A fundamental link between structural network properties and dynamical features of deterministic reaction networks has been known since the 70-80ies with the work of Horn, Jackson and Feinberg [13, 9]. Specifically, their theory concerns the existence and uniqueness of equilibria in *complex balanced* systems, with the 'deficiency zero theorem' playing a central role in this context. Complex balanced systems were called cyclic balanced systems by Boltzmann. They have attractable analytical and physical properties; for example a (pseudo-)entropy might be defined which increases along all trajectories (Boltzmann's H-theorem) [6, 13].

A parallel theory for the stochastic regime is not available, and the very concept of "complex balanced" does not currently have a stochastic counterpart. In this paper we develop a theory to fill this gap. We define *stochastically complex balanced* systems through properties of the stationary distribution, and we prove results for stochastic reaction networks that are in direct correspondence with the results for deterministic models. In particular, we prove a parallel statement of the deficiency zero theorem and show that all deficiency zero reaction network have product-form Poisson-like stationary distributions, irrespectively whether they are complexed balanced or not. In fact, in the non-complexed balanced case, the network is complex balanced on the boundary.

A second target of our study concerns product-form stationary distributions. Such distributions are computationally and analytically tractable and appear in many areas of applied probability, such as, queueing theory [15, 17], Petri Net theory [20], and stochastic reaction network theory [22, 19, 2]. Specifically, a complex balanced mass-action network has a product-form Poisson-like stationary distribution on every irreducible component [19, 2]. As an example, the stationary distribution of (1.1), which is complex balanced, is

$$\pi_\Gamma(x) = M_\Gamma \frac{\kappa_1^{x_A}\kappa_2^{x_B}\kappa_1^{x_C}}{x_A!x_B!x_C!} \quad \text{for } x \in \Gamma,$$

where $\Gamma = \{x \in \mathbb{N}^3 \colon x_A + x_B + 2x_C = \theta\}$ is an irreducible component of the state space $\mathbb{N}^3$ and M_Γ is a normalising constant.

We expand the above result on mass-action systems and give general conditions under which the converse statement is true. In particular, we are interested in providing a structural characterisation of the networks with product-form Poisson-like stationary distributions. However, this class of networks is strictly larger than that of complex balanced networks, and a full characterisation seems hard to achieve. We illustrate this with examples.

2 Background

We first introduce the necessary notation and background material; see [4, 9, 8] for general references. We assume standard knowledge about continuous-time Markov chains.

2.1 Notation

We let $\mathbb{R}$, $\mathbb{R}_0$ and $\mathbb{R}_+$ be the real, the non-negative real and the positive real numbers, respectively. Also let $\mathbb{N}$ be the natural numbers including 0.

For any real number $a \in \mathbb{R}$, $|a|$ denotes the absolute value of a. Moreover, for any vector $v \in \mathbb{R}^p$, we let v_i be the ith component of v, $\|v\|$ the Euclidean norm, and $\|v\|_\infty$ the infinity norm, that is, $\|v\|_\infty = \max_i |v_i|$. For two vectors $v, w \in \mathbb{R}^p$, we write $v < w$ (resp. $v > w$) and $v \leq w$ (resp. $v \geq w$), if the inequality holds component-wise. Further, we define $\mathbb{1}_{\{v \leq w\}}$ to be one if $v \leq w$, and zero otherwise, and similarly for the other inequalities. If $v > 0$ then v is said to be positive. Finally, $\operatorname{supp} v$ denotes the index set of the non-zero components. For example, if $v = (0, 1, 1)$ then $\operatorname{supp} v = \{2, 3\}$.

If $x \in \mathbb{R}_0^q$ and $v \in \mathbb{N}^q$, we define

$$x^v = \prod_{i=1}^q x_i^{v_i}, \quad \text{and} \quad v! = \prod_{i=1}^q v_i!,$$

with the conventions that $0! = 1$ and $0^0 = 1$.

2.2 Reaction networks

A reaction network is a triple $\mathcal{G} = (\mathcal{X}, \mathcal{C}, \mathcal{R})$, where $\mathcal{X} = \{S_1, S_2, \dots, S_n\}$ is a set of n species, $\mathcal{C}$ is a set of m complexes, and $\mathcal{R} \subseteq \mathcal{C} \times \mathcal{C}$ is a set of k reactions, such that $(y, y) \notin \mathcal{R}$ for all $y \in \mathcal{C}$. The complexes are linear combinations of species on $\mathbb{N}$, identified as vectors in $\mathbb{R}^n$. A reaction $(y, y') \in \mathcal{R}$ is denoted by $y \to y'$. We require that every species is part of at least one complex, and that every complex is part of at least one reaction, such that there are no "redundant" species or complexes. In that case, $\mathcal{G}$ is determined by $\mathcal{R}$ (which we allow to be empty). In (1.1), there are $n = 3$ species (A, B, C), $m = 2$ complexes $(A + B, 2B)$, and $k = 2$ reactions.

Given a reaction network $\mathcal{G}$, the *reaction graph* of $\mathcal{G}$ is the directed graph with node set $\mathcal{C}$ and edge set $\mathcal{R}$. We let ℓ be the number of linkage classes (connected components) of the reaction graph. A reaction $y \to y' \in \mathcal{R}$ is *terminal* if any directed path that starts with $y \to y'$ is contained in a closed directed path. We let $\mathcal{R}^*$ be the set of terminal reactions.

A reaction network $\mathcal{G}$ is *weakly reversible*, if every reaction is terminal. The network in (1.1) is weakly reversible, since both reactions are terminal.

The *stoichiometric subspace* of $\mathcal{G}$ is the linear subspace of $\mathbb{R}^n$ given by

$$S = \operatorname{span}(y' - y | y \to y' \in \mathcal{R}).$$

For $v \in \mathbb{R}^n$, the sets $(v + S) \cap \mathbb{R}^n_0$ are called the *stoichiometric compatibility classes* of $\mathcal{G}$ (Fig. 1*A*). For the network in (1.1), $S = \operatorname{span}((-1, 1, 0), (0, 1, -1)) \subset \mathbb{R}^3$, which is 2-dimensional.

2.3 Dynamical systems

We will consider a reaction network $\mathcal{G}$ either as a deterministic dynamical system on the continuous space $\mathbb{R}^n_0$, or as a stochastic dynamical system on the discrete space $\mathbb{N}^n$.

In the deterministic case, the evolution of the species concentrations $z = z(t) \in \mathbb{R}^n_0$ at time t is modelled as the solution to the ODE

$$\frac{dz}{dt} = \sum_{y \to y' \in \mathcal{R}} (y' - y) \lambda_{y \to y'}(z), \tag{2.1}$$

for some functions $\lambda_{y \to y'} \colon \mathbb{R}^n_0 \to \mathbb{R}_0$ and an initial condition $z(0) \in \mathbb{R}^n_0$. We require that the functions $\lambda_{y \to y'}$ are continuously differentiable, and that $\lambda_{y \to y'}(z) > 0$ if and only if $\operatorname{supp} y \subseteq \operatorname{supp} z$. Such functions are called *rate functions*, they constitute a *deterministic kinetics* K for $\mathcal{G}$, and the pair $(\mathcal{G}, K)$ is called a *deterministic reaction system*. If $\lambda_{y \to y'}(z) = \kappa_{y \to y'} z^y$ for all reactions, then the constants $\kappa_{y \to y'}$ are referred to as *rate constants* and the modelling regime is referred to as *deterministic mass-action kinetics*. In this case, the pair $(\mathcal{G}, \kappa)$ is called a *deterministic mass-action system*, where $\kappa \in \mathbb{R}^k_+$ is the vector of rate constants.

In the stochastic setting, the evolution of the species counts $X(t) \in \mathbb{N}^n$ at time t is modelled as a continuous-time Markov chain with state space $\mathbb{N}^n$. At any state $x \in \mathbb{N}^n$, the states that can be reached in one step are $x + y' - y$ for $y \to y' \in \mathcal{R}$, with transition rates $\lambda_{y \to y'}(x)$. The functions $\lambda_{y \to y'} \colon \mathbb{N}^n \to \mathbb{R}_0$ are called *rate functions*, and we require that $\lambda_{y \to y'}(x) > 0$ if and only if $x \geq y$. A choice of these functions constitute a *stochastic kinetics* K for $\mathcal{G}$ and the pair $(\mathcal{G}, K)$ is called a *stochastic reaction system*. If the reaction $y \to y'$ occurs at time t, then the new state is

$$X(t) = X(t-) + y' - y,$$

where $X(t-)$ denotes the previous state. If for any reaction $y \to y' \in \mathcal{R}$

$$\lambda_{y \to y'}(x) = \kappa_{y \to y'} \frac{x!}{(x - y)!} \mathbb{1}_{\{x \geq y\}}.$$

then the constants $\kappa_{y \to y'}$ are known as *rate constants*, as in the deterministic case, and the modelling regime is referred to as *stochastic mass-action kinetics*. The pair $(\mathcal{G}, \kappa)$ is, in this case, called a *stochastic mass-action system*.

The evolution of the stochastic as well as the deterministic reaction system is confined to the stoichiometric compatibility classes,

$$z(t) \in (z(0) + S) \cap \mathbb{R}^n_0 \quad \text{and} \quad X(t) \in (X(0) + S) \cap \mathbb{R}^n_0.$$

In fact, $X(t) \in (X(0)+S) \cap \mathbb{N}^n$, as $X(t)$ takes values in $\mathbb{N}^n$.

Definition 1. *Let $\mathcal{G} = (\mathcal{X}, \mathcal{C}, \mathcal{R})$ be a reaction network.*

a) A reaction network $\mathcal{G}' = (\mathcal{X}', \mathcal{C}', \mathcal{R}')$ is a subnetwork *of $\mathcal{G}$ if $\mathcal{R}' \subseteq \mathcal{R}$.*

b) A system $(\mathcal{G}', K')$, deterministic or stochastic, is a subsystem *of a system $(\mathcal{G}, K)$ if $\mathcal{G}'$ is a subnetwork of $\mathcal{G}$ and the rate functions agree on the reactions in $\mathcal{R}'$.*

c) The subnetwork $\mathcal{G}^* = (\mathcal{X}^*, \mathcal{C}^*, \mathcal{R}^*)$ is the terminal network *of $\mathcal{G}$. Furthermore, the subsystem $(\mathcal{G}^*, K^*)$ of $(\mathcal{G}, K)$ is called the* terminal system *of $(\mathcal{G}, K)$.*

Definition 2. *The connected components of the reaction graph of the terminal network of $\mathcal{G}$ are called* terminal strongly connected component *of $\mathcal{G}$. For any complex y in $\mathcal{C}^*$, we denote by $(\mathcal{G}_y, K_y)$ the subsystem of $\mathcal{G}$ whose reaction graph is the terminal strongly connected component containing y as node.*

As an example, consider the mass-action system

$$2A \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} 2B \xleftarrow{\kappa_3} A \xrightarrow{\kappa_4} 0 \underset{\kappa_6}{\overset{\kappa_5}{\rightleftharpoons}} C.$$

Here, there are two terminal strongly connected components, which are $2A \rightleftharpoons 2B$ and $0 \rightleftharpoons C$. In particular, $(\mathcal{G}_{2A}, K_{2A})$ is equal to $(\mathcal{G}_{2B}, K_{2B})$ and is given by

$$2A \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} 2B.$$

Finally, if $(\mathcal{G}, \kappa)$ is a mass-action system, any subsystems $(\mathcal{G}', K')$ is a mass-action systems as well and can be denoted by $(\mathcal{G}', \kappa')$.

3 Deterministic reaction systems

In this section we will recapitulate the known characterisation of existence and uniqueness of positive equilibria in complex balanced systems and the connection between complex balanced systems and deficiency zero reaction networks. As we will show in the subsequent section, this characterisation can be fully translated into a similar characterisation for stochastic reaction networks.

3.1 Complex balanced systems

We start with a definition.

Definition 3. *A deterministic reaction system $(\mathcal{G}, K)$ is said to be* complex balanced *if there exists a* complex balanced equilibrium*, that is, a positive equilibrium point $c \in \mathbb{R}^n_+$ for the system (2.1), such that*

$$\sum_{y' \in \mathcal{C}} \lambda_{y \to y'}(c) = \sum_{y' \in \mathcal{C}} \lambda_{y' \to y}(c) \quad \textit{for all } y \in \mathcal{C}. \tag{3.1}$$

The name 'complex balanced' refers to the fact that the flow, at equilibrium, entering into the complex y equals the flow exiting from the complex. As an example, the mass-action system in (1.1) is complex balanced and $c = (\kappa_2, \kappa_1, \kappa_1)$ is a complex balanced equilibrium. The class of complex balanced systems is an extension of the class of detailed balanced mass-action systems [13, 9].

For mass-action systems, (3.1) becomes

$$\sum_{y' \in \mathcal{C}} \kappa_{y \to y'} c^y = \sum_{y' \in \mathcal{C}} \kappa_{y' \to y} c^{y'} \quad \text{for all } y \in \mathcal{C}, \tag{3.2}$$

with the convention that $k_{y \to y'} = 0$ if $y \to y' \notin \mathcal{R}$.

In the case of mass-action kinetics, we extend Definition 3 to the stochastic case, by saying that a stochastic mass-action system $(\mathcal{G}, \kappa)$ is complex balanced if the deterministic mass-action system $(\mathcal{G}, \kappa)$ is complex balanced. We might therefore refer to complex balanced mass-action systems without specifying whether they are stochastically or deterministically modelled.

The next theorem is a classical result [13], which provides the backbone for the further characterisation:

Theorem 3.1. *If a deterministic reaction system $(\mathcal{G}, K)$ is complex balanced, then $\mathcal{G}$ is weakly reversible. Moreover, if K is mass-action kinetics, all positive equilibria are complex balanced, that is, fulfil* (3.2)*. Moreover, there exists exactly one complex balanced equilibrium in each stoichiometric compatibility class, which is locally asymptotically stable.*

3.2 Deficiency zero statements

The *deficiency* plays an important role in the study of complex balanced systems. The deficiency of $\mathcal{G}$ is defined as

$$\delta = m - \ell - s,$$

where s is the dimension of the stoichiometric subspace S [13]. The definition hides the geometrical interpretation of the deficiency, which we now will explore.

Let $\{e_y\}_{y\in\mathcal{C}}$ be a basis of $\mathbb{R}^m$. Further, define

$$d_{y\to y'} = e_{y'} - e_y \quad \text{and} \quad \xi_{y\to y'} = y' - y$$

for $y \to y' \in \mathcal{R}$. Let $D = \operatorname{span}(d_{y\to y'} | y \to y' \in \mathcal{R})$. Then $\dim D = m - \ell$ [13].

The space D is linearly isomorphic to the stoichiometric subspace S if and only if $\delta = 0$. Specifically, consider the homomorphism

$$\begin{array}{rccc} \varphi\colon & \mathbb{R}^{|\mathcal{C}|} & \to & \mathbb{R}^n \\ & e_y & \mapsto & y. \end{array} \tag{3.3}$$

For $y \to y' \in \mathcal{R}$, we have $\varphi(d_{y\to y'}) = \xi_{y\to y'}$ and $\varphi_{|D}\colon D \to S$ is thus a surjective homomorfism. Therefore,

$$\delta = \dim D - s = \dim \operatorname{Ker} \varphi_{|D}, \tag{3.4}$$

which implies that $\varphi_{|D}$ is an isomorphism if and only if $\delta = 0$. It further follows that the deficiency is a non-negative number.

We state here a useful Lemma on the deficiency of subnetworks.

Lemma 3.2. *Let $\mathcal{G}$ be a reaction network with deficiency δ. Then, the deficiency of any subnetwork of $\mathcal{G}$ is smaller than or equal to δ.*

Proof. Let $\mathcal{R}' \subseteq \mathcal{R}$ and let $\mathcal{G}'$ be the corresponding subnetwork with deficiency δ'. Further, let D' and S' be the equivalent of D and S for $\mathcal{G}'$, respectively. By (3.4) and since D' is a subspace of D, we have $\delta' = \dim \operatorname{Ker} \varphi_{|D'} \leq \dim \operatorname{Ker} \varphi_{|D} = \delta$, which concludes the proof. □

We next state two classical results which elucidate the connection between complex balanced systems and deficiency zero systems [13, 9]. In particular, the second draws a connection between graphical and dynamical properties of a network. It is given here in a wider formulation than in [9] (see Appendix B for a proof).

Theorem 3.3. *The mass-action system $(\mathcal{G}, \kappa)$ is complex balanced for any choice of κ if and only if $\mathcal{G}$ is weakly reversible and its deficiency is zero.*

Theorem 3.4. *Consider a deterministic reaction system $(\mathcal{G}, K)$, and assume that the deficiency of $\mathcal{G}$ is zero. If $x \in \mathbb{R}^n_0$ is an equilibrium point and $y \to y' \in \mathcal{R}$, then* $\operatorname{supp} y \subseteq \operatorname{supp} x$ *only if $y \to y'$ is terminal. Moreover, if K is mass-action kinetics with rate constants κ and* $\operatorname{supp} y \subseteq \operatorname{supp} x$*, then the projection of x onto the species space of $\mathcal{G}_y$ is a complex balanced equilibrium of $(\mathcal{G}_y, \kappa_y)$.*

It follows from Theorem 3.4 that an equilibrium point satisfies (3.2) for the terminal system, though it is not necessarily a positive equilibrium of $(\mathcal{G}^*, \kappa^*)$ and cannot therefore be considered complex balanced. The deficiency zero theorem, in the following formulation, is a consequence of the three previous theorems:

Theorem 3.5 (Deficiency zero theorem)**.** *Consider a deterministic reaction system $(\mathcal{G}, K)$ for which the deficiency is zero. Then the following statements hold:*

i) if $\mathcal{G}$ is not weakly reversible, then there exists no positive equilibria;

ii) if $\mathcal{G}$ is weakly reversible and K is mass-action kinetics, then there exists within each stoichiometric compatibility class a unique positive equilibrium, which is asymptotically stable.

The original formulation is richer than the one presented here [9].

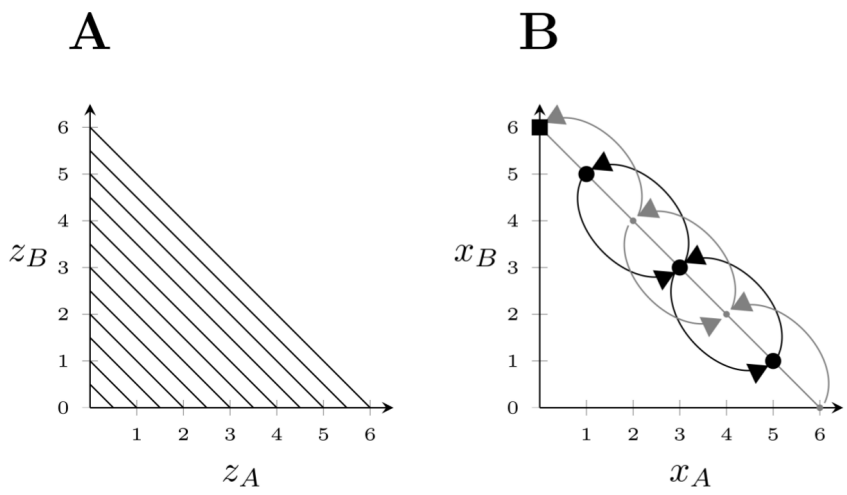


Figure 1: The figure shows some features of the reaction network $2A \to 2B$, and $A + 3B \to 3A + B$. (A) The stoichiometric compatibility classes are of the form $\{(z_A, z_B)\colon z_A + z_B = \text{const.}\}$. (B) The two irreducible components on $\{(x_A, x_B)\colon x_A + x_B = 6\}$ are shown (black circles and square), together with the possible transitions between the states. All states within a component are accessible from each other. The "square" component has no active reactions, both reactions are active on the "black circles" component. The grey states are transient states which are not in any irreducible component.

4 Stochastic reaction systems

4.1 Classification of states and sets

To characterise the stochastic dynamics we introduce the following terminology.

Definition 4. *Let $\mathcal{G} = (\mathcal{X}, \mathcal{C}, \mathcal{R})$ be a reaction network.*

a) A reaction $y \to y' \in \mathcal{R}$ is active *on $x \in \mathbb{N}^n$ if $x \geq y$.*

b) A state $u \in \mathbb{N}^n$ is accessible *from a state $x \in \mathbb{N}^n$ if there is a sequence of $q \geq 0$ reactions $(y_j \to y'_j)_{j=1,\dots,q}$ such that*

(i) $u = x + \sum_{j=1}^{q}(y'_j - y_j)$,

(ii) $y_j \to y'_j$ is active on $x + \sum_{j=1}^{h-1}(y'_j - y_j)$ for all $1 < h \leq q$.

Definition 5. *Let $\mathcal{G}$ be a reaction network. A non-empty set $\Gamma \subseteq \mathbb{N}^n$ is an* irreducible component *of $\mathcal{G}$ if for all $x \in \Gamma$ and all $u \in \mathbb{N}^n$, u is accessible from x if and only if $u \in \Gamma$.*

Definition 6. *A reaction network $\mathcal{G}$ is* essential *if the state space is a union of irreducible components. A reaction network $\mathcal{G}$ is* almost essential *if the state space is a union of irreducible components except for a finite number of states.*

'Irreducible' and 'essential' are standard terms in Markov chain theory. An essential network is also almost essential. A reaction network is essential if and only if every state of the associated Markov chain is 'essential' [11].

A weakly reversible reaction network is essential [21]. Conditions for being essential can be found in [21, 10]. Any irreducible component is contained in some stoichiometric compatibility class, and a stoichiometric compatibility class may contain several irreducible components (Fig. 1B).

4.2 Stationary distribution

The stationary distribution π_Γ on an irreducible component Γ is unique, if it exists. It is characterised by the *master equation* [4]:

$$\sum_{y\to y'\in\mathcal{R}} \pi_\Gamma(x+y-y')\lambda_{y\to y'}(x+y-y') = \pi_\Gamma(x)\sum_{y\to y'\in\mathcal{R}} \lambda_{y\to y'}(x), \tag{4.1}$$

for all $x \in \Gamma$. Let $X(t)$ denote the stochastic process associated with the system. If $X(t_0)$ follows the law of π_Γ at time t_0, then the distribution of $X(t)$ is π_Γ for all future times $t \geq t_0$. In this sense, the stationary distribution describes a state of equilibrium of the system. Moreover, if π_Γ exists, then

$$\lim_{t\to\infty} P(X(t) \in A) = \pi_\Gamma(A) \quad \text{for any } A \subseteq \Gamma, \tag{4.2}$$

provided that $X(0) \in \Gamma$ with probability one. As discussed in Section 1, a connection between mass-action complex balanced systems and their stationary distribution has been made in [2]:

Theorem 4.1. *Let $(\mathcal{G}, \kappa)$ be a complex balanced mass-action system. Then, there exists a unique stationary distribution on every irreducible component Γ, and it is of the form*

$$\pi_\Gamma(x) = M_\Gamma^c \prod_{i=1}^n \frac{c_i^{x_i}}{x_i!} \quad \textit{for } x \in \Gamma, \tag{4.3}$$

where c is a complex balanced equilibrium of $(\mathcal{G}, \kappa)$ and M_Γ^c is a normalising constant.

The result in [2] assumes deficiency zero, but the proof only uses the fact that c is a complex balanced equilibrium, and therefore it holds in the broader setting of complex balanced systems. Our results will expand Theorem 4.1.

4.3 Parallel theorems for stochastic mass-action systems

In this section we derive stochastic statements corresponding to Theorem 3.1-3.5. Some of the proofs are deferred to Appendix B. We begin with a definition.

Definition 7. *For an irreducible component Γ, the set $\mathcal{R}_\Gamma$ of* active reactions *on Γ consists of the reactions $y \to y' \in \mathcal{R}$ that are active on some $x \in \Gamma$. The subnetwork $\mathcal{G}_\Gamma = (\mathcal{X}_\Gamma, \mathcal{C}_\Gamma, \mathcal{R}_\Gamma)$ is called the* Γ-network *of $\mathcal{G}$ and the subsystem $(\mathcal{G}_\Gamma, K_\Gamma)$ of $(\mathcal{G}, K)$ is called the* Γ-system *of $(\mathcal{G}, K)$.*

The reactions that are active on Γ determine the dynamics of the stochastic system on Γ. To study the stationary distributions, it is therefore convenient to analyse the Γ-systems. Note that $\mathcal{R}_\Gamma$ is empty if and only if Γ consists of a single state.

As an example, consider the deficiency zero network,

$$C \rightleftharpoons D, \quad 2A \rightleftharpoons 2B, \quad A \to 0.$$

All molecules of A and B are irreversibly consumed through $A \to 0$ and $2B \to 2A$, thus the only active reactions on an irreducible component $\Gamma \neq \{0\}$ are $C \rightleftharpoons D$. The Γ-network is therefore $C \rightleftharpoons D$, which differs from the terminal system $C \rightleftharpoons D$, $2A \rightleftharpoons 2B$. The next proposition states that for a deficiency zero reaction network $\mathcal{R}_\Gamma \subseteq \mathcal{R}^*$ for any irreducible component Γ. It does not hold in general, for example,

$$A \to B, \quad 2B \to 2A$$

has $\mathcal{R}_\Gamma = \mathcal{R}$ for any $\Gamma \neq \{0\}, \{(0,1)\}$, while $\mathcal{R}^* = \emptyset$.

Proposition 4.2. *Let $\mathcal{G}$ be a reaction network and Γ an irreducible component such that $\mathcal{G}_\Gamma$ has deficiency zero. Then, $\mathcal{G}_\Gamma$ is a subnetwork of $\mathcal{G}^*$. In particular, this is true if the deficiency of $\mathcal{G}$ is zero.*

See Appendix B for a proof. Proposition 4.2 can be useful because $\mathcal{R}_\Gamma$ might be difficult to find, especially if there are many complexes. On the other hand, terminal reactions are easily identified by means of the reaction graph. The next definitions are inspired by Definition 3.

Definition 8. *Let $(\mathcal{G}, K)$ be a stochastic reaction system. A stationary distribution on an irreducible component Γ is said to be* complex balanced *if*

$$\sum_{y\in\mathcal{C}_\Gamma} \pi_\Gamma(x-y'+y)\lambda_{y\to y'}(x-y'+y) = \sum_{y\in\mathcal{C}_\Gamma} \pi_\Gamma(x)\lambda_{y'\to y}(x) \quad \forall y'\in\mathcal{C}_\Gamma, x\in\Gamma. \tag{4.4}$$

For a mass-action system, (4.4) becomes

$$\sum_{y\in\mathcal{C}_\Gamma} \pi_\Gamma(x-y'+y)\kappa_{y\to y'}\frac{(x-y'+y)!}{(x-y')!}\mathbb{1}_{\{x\geq y'\}} = \sum_{y\in\mathcal{C}_\Gamma} \pi_\Gamma(x)\kappa_{y'\to y}\frac{x!}{(x-y')!}\mathbb{1}_{\{x\geq y'\}}$$

for any $y'\in\mathcal{C}_\Gamma$ and $x\in\Gamma$, with the convention that $k_{y\to y'}=0$ if $y\to y'\notin\mathcal{R}_\Gamma$. In developing the theory for complex balanced equilibria in the deterministic setting, an important role is played by the fact that a complex balanced equilibrium is positive by definition. Our aim is to introduce a similar concept for the stochastic systems. In the deterministic setting, if a state $z\in\mathbb{R}^n$ is positive then every rate function calculated on z is positive. We find inspiration from this to give the next definition:

Definition 9. *An irreducible component Γ is* positive *if $\mathcal{G}_\Gamma=\mathcal{G}$.*

Equivalently, an irreducible component Γ is *positive* if all reactions are active on Γ. The next definition follows naturally by analogy with the deterministic setting.

Definition 10. *A stochastic reaction system $(\mathcal{G},\kappa)$ is said to be* stochastically complex balanced *if there exists a complex balanced stationary distribution on a positive irreducible component.*

If Γ is positive, then $\mathcal{C}_\Gamma=\mathcal{C}$ and a complex balanced stationary distribution on Γ satisfies (4.4) with $\mathcal{C}_\Gamma$ replaced by $\mathcal{C}$. Note the similarity between Definition 10 and the definition of a complex balance equilibrium (Definition 3): the positivity of Γ plays the role of the positivity in Definition 3. Also note the close similarity between (3.1) and (4.4).

Theorem 4.3. *Let $(\mathcal{G},K)$ be a stochastic reaction system, and let Γ be an irreducible component. If there exists a complex balanced stationary distribution π_Γ on Γ then $\mathcal{G}_\Gamma$ is weakly reversible. Moreover, if K is mass-action kinetics with rate constants κ, there exists a complex balanced stationary distribution π_Γ on Γ if and only if the Γ-system of $(\mathcal{G},\kappa)$ is complex balanced. If this is case, then π_Γ has the form*

$$\pi_\Gamma(x) = M_\Gamma^c \prod_{i:\ S_i\in\mathcal{X}_\Gamma} \frac{c_i^{x_i}}{x_i!} \quad \text{for } x\in\Gamma, \tag{4.5}$$

where c is a complex balanced equilibrium of $(\mathcal{G}_\Gamma,\kappa_\Gamma)$ and M_Γ^c is a normalising constant.

The proof is in Appendix B. It is shown in [2] that the stationary distribution $\pi_\Gamma(x)$ is independent of the choice of complex balanced equilibrium c of the Γ-system. We are now ready to derive stochastic versions of Theorem 3.1-3.5. In addition, we will show that a stochastically complexed balanced mass-action system is complex balanced and *vice versa*. Hence, we will show that the deterministic and stochastic systems are intimately connected. The next corollary is an equivalent of Theorem 3.1.

Corollary 4.4. *If a stochastic reaction system $(\mathcal{G},K)$ is stochastically complex balanced then $\mathcal{G}$ is weakly reversible. Moreover, a mass-action system $(\mathcal{G},\kappa)$ is stochastically complex balanced if and only if it is complex balanced. If this is case, then on every irreducible component Γ there exists a unique stationary distribution π_Γ. Such π_Γ is a complex balanced stationary distribution and it has the form* (4.3)*, where c is a complex balanced equilibrium of $(\mathcal{G},\kappa)$.*

Proof. If Γ is positive, then $(\mathcal{G}_\Gamma,K_\Gamma)=(\mathcal{G},K)$. Therefore, by Theorem 4.3 if $(\mathcal{G},K)$ is stochastically complex balanced then $\mathcal{G}$ is weakly reversible. Moreover, if K is mass-action kinetics with rate constants κ, it follows from Theorem 4.3 that there exists a complex balanced stationary distribution on Γ if and only if $(\mathcal{G},\kappa)$ is complex balanced. In this case, by Theorem 4.1, a stationary distribution exists on every irreducible component and it is of the form (4.3). By Theorem 4.3, it is a complex balanced stationary distribution. □

Corollary 4.4 might be considered a stochastic version of Theorem 3.1, especially if (4.2) is taken to be equivalent to "asymptotic stability" for a deterministic equilibrium. Part of the corollary is known [2] (see also Theorem 4.1), and the whole corollary might therefore be considered as an extension of the result in [2] on mass-action systems. In this sense, Theorem 4.3 provides an even more general version, which deals with complex balanced subsystems of $(\mathcal{G}, \kappa)$.

We now state the parallel versions of Theorem 3.3-3.5 for the stochastic setting.

Corollary 4.5. *The mass-action system $(\mathcal{G}, \kappa)$ is stochastically complex balanced for any choice of κ if and only if $\mathcal{G}$ is weakly reversible and its deficiency is zero.*

Proof. The result is an immediate consequence of Corollary 4.4 and Theorem 3.3. □

Theorem 4.6. *Consider a stochastic reaction system $(\mathcal{G}, K)$, and assume the deficiency of $\mathcal{G}$ is zero. Let x be a state in an irreducible component Γ and let $y \to y'$ in $\mathcal{R}$. Then, $y \leq x$ only if $y \to y'$ is terminal. Moreover, if K is mass-action kinetics, then on Γ the stationary distribution has the form*

$$\pi_\Gamma(x) = M_\Gamma^c \prod_{i:\, S_i \in \mathcal{X}^*} \frac{c_i^{x_i}}{x_i!} \quad \text{for } x \in \Gamma, \tag{4.6}$$

where c is a complex balanced equilibrium for the terminal system, and M_Γ^c is a normalising constant.

The proof is in Appendix B.

Theorem 4.7. *Consider a stochastic reaction system $(\mathcal{G}, K)$, and assume that the deficiency of $\mathcal{G}$ is zero. Then the following statements hold:*

i) if $\mathcal{G}$ is not weakly reversible, then there exist no positive irreducible components;

ii) if $\mathcal{G}$ is weakly reversible, then $\mathcal{G}$ is essential, and if K is mass-action kinetics there exists a unique stationary distribution on every irreducible component.

The proof of the theorem is in Appendix B. In case (i), Theorem 4.6 provides the form of the stationary distribution. Hence we have characterised the stationary distribution for any deficiency zero reaction system, irrespectively whether it is complex balanced or not.

Example 1. Consider the two stochastic mass-action systems

$$A \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} B, \quad 10A \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} 10B \quad \text{and} \quad A \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} B, \quad 10A \xrightarrow{\kappa_3} 0.$$

The behaviours of the two corresponding deterministic systems differ substantially, while the behaviours of the stochastic systems are equivalent on the irreducible components $\Gamma_\theta = \{x \in \mathbb{N}^2 \colon x_1 + x_2 = \theta\}$ with $0 \leq \theta < 10$ an integer. Indeed, in both cases the Γ_θ-system is

$$A \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} B,$$

which is complex balanced (Theorem 3.3). It follows from Theorem 4.3 that the stationary distribution on Γ_θ is

$$\pi_\theta(x_1, x_2) = M_\theta \frac{\kappa_2^{x_1}}{x_1!} \frac{\kappa_1^{x_2}}{x_2!} \quad \text{for } (x_1, x_2) \in \Gamma_\theta,$$

for a suitable normalizing constant M_θ. The stationary distributions is complex balanced, but since Γ_θ is not positive in either of the two networks, we cannot conclude that the systems are stochastically complex balanced. Indeed, they are not for some choice of rate constants (Corollary 4.5).

5 Product-form Poisson-like stationary distributions

The above results draw parallels between stochastic and deterministic reaction networks. If a mass-action system is (stochastically) complex balanced, then the stationary distribution on every irreducible component is a product-form Poisson-like distribution. Does the reverse statement hold true too? If the stationary distribution is a product-form Poisson-like distribution on some, or all irreducible components, does it follow that the system is complex balanced? In the spirit of the first part of the paper we would like to achieve a full characterisation of stochastic systems with product-form Poisson-like stationary distributions. However, even though the hypothesis of Theorem 5.1 below is rather general, a full characterisation seems hard to achieve.

Theorem 5.1. *Let $\mathcal{G}$ be an almost essential reaction network, $\kappa \in \mathbb{R}^m_+$ a vector of rate constants and $c \in \mathbb{R}^n_+$ a vector with positive entries. The probability distribution $\pi_\Gamma \colon \Gamma \to (0,1]$, defined by (4.3) is a stationary distribution for the stochastic mass-action system $(\mathcal{G}, \kappa)$ for all irreducible components $\Gamma \subseteq \mathbb{N}^n$ of $\mathcal{G}$ if and only if c is a complex balanced equilibrium for $(\mathcal{G}, \kappa)$.*

Proof. By Theorem 4.1, if c is a complex balanced equilibrium for $(\mathcal{G}, \kappa)$, then the stationary distribution on all irreducible components $\Gamma \subseteq \mathbb{N}^n$ is of the form (4.3).

Oppositely, assume that (4.3) is the stationary distribution on Γ for the stochastic mass-action system $(\mathcal{G}, \kappa)$, for all irreducible components Γ. Since $\mathcal{G}$ is almost essential, there exists a constant K such that any states x with $\|x\| > K$ belongs to an irreducible component Γ. For any $x \in \mathbb{N}^n$, such that

$$\min_{S_i \in \mathcal{X}} x_i > \max_{y \to y' \in \mathcal{R}} (\|y\|_\infty + \|y'\|_\infty) + K, \tag{5.1}$$

we have that $x \geq y$ and $x - y' + y \geq y$ for all $y \to y' \in \mathcal{R}$. Then, since (4.3) is a stationary distribution and since x and $x + y - y'$ are in the same irreducible component for all $y \to y' \in \mathcal{R}$, we have from (4.1)

$$\sum_{y \to y' \in \mathcal{R}} \pi_\Gamma(x + y - y') \kappa_{y \to y'} \frac{(x + y - y')!}{(x - y')!} = \pi_\Gamma(x) \sum_{y \to y' \in \mathcal{R}} \kappa_{y \to y'} \frac{x!}{(x - y)!}, \tag{5.2}$$

for all $x \in \Gamma$ satisfying (5.1). Further, using (4.3), equation (5.2) becomes

$$\sum_{y \to y' \in \mathcal{R}} \frac{x!}{(x - y')!} \kappa_{y \to y'} c^{y - y'} = \sum_{y \to y' \in \mathcal{R}} \kappa_{y \to y'} \frac{x!}{(x - y)!},$$

which, by rearranging terms, leads to

$$\sum_{y' \in \mathcal{C}} \frac{x!}{(x - y')!} \sum_{y \to y' \in \mathcal{R}} \kappa_{y \to y'} c^{y - y'} = \sum_{y' \in \mathcal{C}} \frac{x!}{(x - y')!} \sum_{y' \to y \in \mathcal{R}} \kappa_{y' \to y}. \tag{5.3}$$

The equality holds for all $x \in \mathbb{N}^n$ satisfying (5.1), therefore the polynomials on the two sides of (5.3) are equal.

For any $y' \in \mathcal{C}$, let $p_{y'}(x)$ be the polynomial

$$p_{y'}(x) = \frac{x!}{(x - y')!}.$$

The monomial with maximal degree in $p_{y'}$ is $x^{y'}$, and these differ for all complexes $y' \in \mathcal{C}$. This implies that $p_{y'}$, $y' \in \mathcal{C}$, are linearly independent on $\mathbb{R}$, and thus, the polynomials on the two sides of (5.3) are equal if and only if

$$\sum_{y \in \mathcal{C}} \kappa_{y \to y'} c^{y - y'} = \sum_{y \in \mathcal{C}} \kappa_{y' \to y} \quad \text{for all } y' \in \mathcal{C}.$$

Hence, c is a complex balanced equilibrium for $(\mathcal{G}, \kappa)$ and the proof is completed. □

5.1 Relaxation of Assumptions in Theorem 5.1

To infer the existence of complex balanced equilibria in Theorem 5.1, the assumptions of the theorem could be weakened. Specifically, it is only required that (5.3) holds for a set of states whose geometry and cardinality allow us to conclude that the polynomials on the two sides of (5.3) are the same. For (5.3) to hold, we need x to be in a irreducible component and we require $x \geq y$ and $x - y' + y \geq y$ for all reactions $y \to y' \in \mathcal{R}$, as well as the stationary distribution evaluated in x and $x - y' + y$ to be of the form (4.3). If a state x satisfies this, we call it a *good state*.

A more general condition than being almost essential could be chosen case by case and depends on the monomials appearing in (5.3). For example, if the set of complexes coincides with the set of species, then the polynomials in (5.3) are linear and the existence of $n+1$ *good states* in general position implies the existence of a complex balanced equilibrium. In general, let d be the total degree of the polynomials in (5.3). Then it is sufficient to have an n-dimensional cube with more than $d+1$ good states on all edges. Therefore, to conclude that a system is complex balanced it is sufficient to check the behaviour of a finite number of states, lying on a finite number of irreducible components. However, it follows from Examples 2 and 4 that the existence of arbitrarily many good states on a few irreducible components does not imply the existence of a complex balanced equilibrium in general. Finally, in order to postulate that the mass-action system is complex balanced, it is necessary that the vector c appearing in Theorem 5.1 is the same for every irreducible component, as shown in Example 5.

The following examples are also meant to give an idea of why it is hard to obtain a full characterization of stochastic mass-action systems with a product-form Poisson-like stationary distribution on some irreducible component.

Example 2. Let $\rho \in \mathbb{R}_+$ and let $\theta \geq 2$ be an integer. Consider the stochastic mass-action system

$$A \xrightarrow{\rho(\theta-1)} B \qquad 2B \xrightarrow{\rho} 2A, \tag{5.4}$$

where $\kappa_1 = \rho(\theta - 1)$ and $\kappa_2 = \rho$ are the rate constants. The reaction network is almost essential. By the master eaquation, it can be shown that the stationary distribution on the irreducible component $\Gamma_\theta = \{x \in \mathbb{N}^2 \colon x_1 + x_2 = \theta\}$ has the form (4.3) with $c = (1, 1)$, namely

$$\pi_\theta(x_1, x_2) = M_\theta \frac{1}{x_1! x_2!} \quad \text{for } (x_1, x_2) \in \Gamma_\theta, \tag{5.5}$$

where M_θ is a normalising constant. However, the mass-action system is not complex balanced as the reaction network is not weakly reversible (Theorem 3.1). In particular, by Theorem 5.1, not all irreducible components can have a stationary distribution of the form (4.3) with $c = (1, 1)$. Trivially, the absorbing states $(0, 0)$ and $(0, 1)$ have it.

Additionally, we should point out that there is not an *equivalent* system on Γ (that is, a stochastic mass-action system with the same transition rate matrix on the states of Γ as (5.4)) which is complex balanced. Consider the case $\theta = 1$. Since the transition from $(0, 2)$ to $(2, 0)$ is possible according to (5.4), any equivalent mass-action system must contain the reaction $2B \to 2A$, with rate constant ρ. It can be further shown that any equivalent weakly reversible mass-action system must contain the connected component

$$A + B \underset{\frac{\rho}{2}}{\overset{\rho}{\rightleftarrows}} 2B \overset{\rho}{\rightarrow} 2A\,.$$

This prevents the system from being complex balanced, since there is not a $c \in \mathcal{R}_+^2$ fulfilling (3.2) for the three complexes $2B$, $2A$ and $A + B$.

Example 3. Let $\rho_1, \rho_2, \rho_3 \in \mathbb{R}_+$ and let $\theta \geq 2$ be an integer. Consider the modification of Example 2 given by

$$A \underset{\rho_2}{\overset{\rho_1(\theta-1)+\rho_2}{\rightleftharpoons}} B \qquad 2B \underset{\rho_3}{\overset{\rho_1+\rho_3}{\rightleftharpoons}} 2A,$$

which is weakly reversible. If we let $\rho_2 = 0$ and $\rho_3 = 0$, then the system reduces to that of Example 2 by removing the two reversible reactions. It can be shown that for any parameter choice, (5.5) is still

a stationary distribution on the irreducible component $\Gamma_\theta = \{x \in \mathbb{N}^2 : x_1 + x_2 = \theta\}$. However, the mass-action system is not complex balanced for any choice of the parameters. It can be further shown that irreducible components different from Γ do not possess a product-form Poisson-like stationary distribution.

Example 4. Consider the stochastic mass-action system with $\rho \in \mathbb{R}_+$ and θ_1, θ_2 two positive integers,

$$\begin{aligned} A &\xrightarrow{\rho\theta_1\theta_2} B & \qquad 2B &\xrightarrow{\rho(\theta_1+\theta_2-1)} 2A \\ 3A &\xrightarrow{\rho} A + 2B & \qquad 2A + B &\xrightarrow{\rho} 3B. \end{aligned}$$

The reaction network is almost essential. For any $\theta \in \mathbb{N}$, consider the irreducible component $\Gamma_\theta = \{x \in \mathbb{N}^2 \colon x_1 + x_2 = \theta + 1\}$. Then π_{θ_1} and π_{θ_2}, defined as in (5.5), are the (unique) stationary distributions on the irreducible components Γ_{θ_1} and Γ_{θ_2}, respectively. However, the mass-action system is not complex balanced, since the reaction network is not weakly reversible (Theorem 3.1).

Example 5. Theorem 5.1 can be also used to compute the stationary distribution of a stochastic mass-action system which behaves as a complex balanced system on the irreducible components. Consider the weakly reversible (and therefore essential) stochastic mass-action system

$$A \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} 2A \qquad A + B \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} 2A + B.$$

On every irreducible component $\Gamma_\theta = \{x \in \mathbb{N}^2 : x_2 = \theta\}$, $\theta \in \mathbb{N}$, the associated continuous time Markov chain, which describes the evolution of the counts of A, has the same distribution as the process associated with

$$A \underset{\kappa_2+\kappa_4\theta}{\overset{\kappa_1+\kappa_3\theta}{\rightleftharpoons}} 2A,$$

because the transition rates coincide. The latter system is complex balanced for any choice of rate constants. The stationary distribution has the form (Theorem 5.1)

$$\pi_\theta(x) = M_\theta \frac{1}{x!} \left(\frac{\kappa_2 + \kappa_4\theta}{\kappa_1 + \kappa_3\theta} \right)^x$$

for some positive constant M_θ. The latter gives the stationary distribution of the original system as well. However, the rate of the Poisson distribution does depend on θ, in which case the original system cannot be complex balanced (Corollary 4.4). For the same reason the example does not contradict Theorem 5.1.

6 Applications

There are not many means to explicitly calculate the stationary distribution of a stochastic mass-action system. As an example, Theorem 4.3 can be used to determine the stationary distributions of mass-action systems like

$$C \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} D, \quad 2A \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} 2B, \quad A \xrightarrow{\kappa_5} 0.$$

Indeed, for any irreducible component Γ different from $\{0\}$, the Γ-system is given by

$$C \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} D,$$

which is weakly reversible and has deficiency zero, therefore it is complex balanced. Hence, the stationary distribution on Γ has the form

$$\pi_\Gamma(x) = M_\Gamma \frac{\kappa_2^{x_3}}{x_3!} \frac{\kappa_1^{x_4}}{x_4!} \quad \text{for } x \in \Gamma,$$

where x_3 and x_4 denote the entries relative to C and D, respectively. Alternatively, since the terminal system is given by

$$C \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} D, \quad 2A \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} 2B,$$

Theorem 4.6 can be used to compute the stationary distribution. On every irreducible component Γ, it is given by

$$\pi_\Gamma(x) = \widetilde{M}_\Gamma \frac{(\sqrt{\kappa_4})^{x_1}}{x_1!} \frac{(\sqrt{\kappa_3})^{x_2}}{x_2!} \frac{\kappa_2^{x_3}}{x_3!} \frac{\kappa_1^{x_4}}{x_4!} \quad \text{for} \quad x \in \Gamma,$$

which is equivalent to the previous formula since x_1 and x_2 are constantly 0 on all irreducible components.

If the system does not fulfil the conditions of Theorem 4.3 and neither can be cast as a birth-death process, Theorem 5.1 might be useful. The following mass-action system is considered in [1]:

$$A \xrightarrow{\kappa_1} 0 \qquad 0 \xrightarrow{\kappa_2} 2A.$$

By Theorem 5.1, the stationary distribution cannot be Poisson. Indeed, it is given by the distribution of $Y = Y_1 + 2Y_2$, where Y_1 and Y_2 are two independent Poisson random variables with rates $\frac{\kappa_2}{\kappa_1}$ and $\frac{\kappa_2}{2\kappa_1}$, respectively. Hence,

$$\pi(x) = e^{-\frac{3\kappa_2}{2\kappa_1}} \sum_{\substack{i,j \in \mathbb{N} \\ x = i+2j}} \frac{1}{i!j!} \left(\frac{\kappa_2}{\kappa_1}\right)^i \left(\frac{\kappa_2}{2\kappa_1}\right)^j.$$

In [1], the following system is also considered:

$$0 \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} A \qquad 2A \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} 3A.$$

It has the stationary distribution

$$\pi(x) = M \prod_{i=1}^{x} \frac{\theta_1[(i-1)(i-2) + \theta_2]}{i(i-1)(i-2) + \theta_3 i} \quad \text{for } x \in \mathbb{N},$$

where $\theta_1 = \kappa_3/\kappa_4$, $\theta_2 = \kappa_1/\kappa_3$, $\theta_3 = \kappa_2/\kappa_4$ and $M = \pi(0)$ is a normalising constant. It is interesting that $\pi(x)$ is a Poisson distribution if and only if $\theta_2 = \theta_3$. In fact, and in accordance with our results, the mass-action system is complex balanced if and only if $\theta_2 = \theta_3$.

7 Discussion

Corollary 4.5 provides a characterisation of reaction networks that are stochastically complex balanced for any choice of rate constants. It is natural to wonder whether a stationary distribution of the form (4.3) on some irreducible component Γ for all choices of rate constants implies something specific about the Γ-system. If for specific form we intend deficiency zero and weakly reversible, this is not the case, as this is violated in Example 5. However, in Example 5 the system might be described *equivalently* by means of a weakly reversible deficiency zero system for any irreducible component. The question of whether this is always true remains open. We provide here two more examples.

Example 6. Consider the stochastic mass-action system

$$2A \xrightarrow{\kappa_1} 2B \qquad A + 3B \xrightarrow{\kappa_2} 3A + B.$$

The underlying reaction network is the one considered in Figure 1. On the irreducible component $\Gamma = \{(1,5), (3,3), (5,1)\}$, the Markov chain associated with the system has the same distribution as the Markov chain associated with

$$2A \underset{3\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} 2B,$$

since the transition rates coincide. It is interesting to note that the dynamics of the two systems are different when they are deterministically modelled [7]. Due to Theorem 3.3, the latter system is complex balanced for any choice of rate constants. Therefore, by Theorem 5.1, the stationary distribution on Γ has the form (4.3) on both systems for any choice of rate constants. The same argument does not hold, in this case, for the other irreducible components.

Example 7. The same phenomenon as in Example 6 is observed in the stochastic mass-action system

$$2A \xrightarrow{\kappa_1} 3A + B \qquad A + 3B \xrightarrow{\kappa_2} 2B.$$

On the irreducible component $\Gamma = \{(x_1, x_2) \in \mathbb{N}^2 : x_1 \geq 2, x_1 = x_2\}$, the Markov chain associated with the system has the same distribution as the Markov chain associated with

$$2A \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} 3A + B,$$

since the transition rates coincide, and the latter network is weakly reversible and has deficiency zero.

Acknowledgments: We thank Elisenda Feliu for valuable comments and fruitful discussions.

A Preliminary results

Here we state some preliminary results that will be needed in Appendix B.

Lemma A.1. *Let $\mathcal{G}$ be a reaction network. If $y_1 \to y_2 \to \cdots \to y_q$ is a directed path in the reaction graph of $\mathcal{G}$, and $x \geq y_1$, then $x + y_q - y_1$ is accessible from x.*

Proof. First, note that

$$x + \sum_{i=1}^{q-1} (y_{i+1} - y_i) = x + y_q - y_1.$$

It is sufficient to note that if $x \geq y_1$, then for any $1 \leq j \leq q-1$, we have

$$x + \sum_{i=1}^{j-1} (y_{i+1} - y_i) = x + y_j - y_1 \geq y_j.$$

This concludes the proof. □

Lemma A.2. *Let Γ be an irreducible component such that $\mathcal{G}_\Gamma$ has deficiency zero. Then, $\mathcal{G}_\Gamma$ is weakly reversible. In particular, if $\mathcal{G}$ has deficiency zero, $\mathcal{G}_\Gamma$ has deficiency zero and is weakly reversible for every irreducible component Γ.*

Proof. If $\mathcal{R}_\Gamma$ is empty then $\mathcal{G}_\Gamma$ is weakly reversible and there is nothing to prove. Otherwise, if $\mathcal{R}_\Gamma$ is non-empty, let $y_1 \to y_1' \in \mathcal{R}_\Gamma$. By hypothesis, there exists a state x in Γ with $x \geq y_1$. This means that $x + \xi_{y_1 \to y_1'}$ is accessible from x. Moreover, since x belongs to an irreducible component Γ, we have that x is accessible from $x + \xi_{y_1 \to y_1'}$ as well, which implies that

$$x = x + \sum_{j=1}^{q} \xi_{y_j \to y_j'},$$

for a certain choice of $\xi_{y_j \to y_j'}$. In particular, $\sum_{j=1}^q \xi_{y_j \to y_j'} = 0$. This implies that $\sum_{j=1}^q d_{y_j \to y_j'} = 0$, because φ, defined in (3.3), is an isomorphism between the spaces D and S associated with $\mathcal{G}_\Gamma$. Therefore,

$$\sum_{j=1}^{q} (e_{y_j'} - e_{y_j}) = \sum_{y \in \mathcal{C}_\Gamma} \alpha_y e_y = 0,$$

for some integers α_y. Since the vectors e_y are linearly independent, $\alpha_y = 0$ for all $y \in \mathcal{C}_\Gamma$. Hence, each e_y that appears in the sum, must appear at least twice, once with coefficient 1, once with -1. Consequently, by iteratively reordering the terms $d_{y_j \to y_j'}$, the reactions $(y_j \to y_j')_{j=1}^q$ form a union of directed closed paths in the reaction graph of $\mathcal{G}$. In particular, the reaction $y_1 \to y_1'$ is contained in a closed directed path of the reaction graph of $\mathcal{G}_\Gamma$, and since this is true for every reaction in $\mathcal{R}_\Gamma$, $\mathcal{G}_\Gamma$ is weakly reversible. We conclude the proof by Lemma 3.2, since if $\mathcal{G}$ has deficiency zero, so does every subnetwork of $\mathcal{G}$. □

Lemma A.3. *Let $\mathcal{G}$ be a weakly reversible reaction network, and let Γ be an irreducible component. Then, for any complex $y' \in \mathcal{C}_\Gamma$ we have*

$$\{y \in \mathcal{C} \colon y \to y' \in \mathcal{R}\} = \{y \in \mathcal{C}_\Gamma \colon y \to y' \in \mathcal{R}_\Gamma\},$$
$$\{y \in \mathcal{C} \colon y' \to y \in \mathcal{R}\} = \{y \in \mathcal{C}_\Gamma \colon y' \to y \in \mathcal{R}_\Gamma\}.$$

Proof. One inclusion is trivial, since $\mathcal{R}_\Gamma \subseteq \mathcal{R}$. For the other inclusion, fix $y' \in \mathcal{C}_\Gamma$. Suppose that there exists $x \in \Gamma$ with $x \geq y'$. It follows that any reaction $y' \to y \in \mathcal{R}$ is active on Γ, and therefore is contained in $\mathcal{R}_\Gamma$. Moreover, since $\mathcal{G}$ is weakly reversible, for any reaction in $\mathcal{R}$ of the form $y \to y'$, there exists a directed path in the reaction graph of $\mathcal{G}$ from y' to y. Hence, by Lemma A.1, $x + y - y'$ is accessible from x, which implies that $x + y - y'$ is in Γ and that $y \to y'$ is in $\mathcal{R}_\Gamma$, since $x + y - y' \geq y$. Therefore, to conclude the proof it suffices to prove that there exists $x \in \Gamma$ with $x \geq y'$.

If it were no $x \in \Gamma$ with $x \geq y'$, then no reaction of the form $y' \to y$ would be in $\mathcal{R}_\Gamma$. Since $y' \in \mathcal{C}_\Gamma$, there exists a reaction of the form $y \to y'$. This means that there is $\tilde{x} \in \Gamma$, such that $\tilde{x} \geq y$. Hence, $\tilde{x} + y' - y$ is in Γ with $\tilde{x} + y' - y \geq y'$, which concludes the proof. □

B Proofs

B.1 Proof of Theorem 3.4

By [9, Theorem 6.1.2], if $x \in \mathbb{R}^n_0$ is an equilibrium point and $y \to y' \in \mathcal{R}$, then $\operatorname{supp} y \subseteq \operatorname{supp} x$ only if $y \to y'$ is terminal. Moreover, if $\operatorname{supp} y \subseteq \operatorname{supp} x$, then $\operatorname{supp} \widetilde{y} \subseteq \operatorname{supp} x$ for every complex $\widetilde{y}$ of $\mathcal{G}_y = (\mathcal{X}_y, \mathcal{C}_y, \mathcal{R}_y)$.

Now, suppose that K is mass-action kinetics with rate constants κ, and that $\operatorname{supp} y \subseteq \operatorname{supp} x$ with $y \to y' \in \mathcal{R}$ (and therefore $y \to y' \in \mathcal{R}^*$). Consider

$$\widetilde{\mathcal{R}} = \{\widetilde{y} \to \widetilde{y}' \in \mathcal{R} \colon \operatorname{supp} \widetilde{y} \subseteq \operatorname{supp} x\}.$$

By the first part of the statement, the reaction graph of the subnetwork $\widetilde{\mathcal{G}} = (\widetilde{\mathcal{X}}, \widetilde{\mathcal{C}}, \widetilde{\mathcal{R}})$ is a union of terminal strongly connected components of $\mathcal{G}$, and therefore $\widetilde{\mathcal{G}}$ is weakly reversible. Moreover, by Lemma 3.2, the deficiency of $\widetilde{\mathcal{G}}$ is 0. It is not hard to see that the canonical projection of x onto the space of the species $\widetilde{\mathcal{X}}$ is a positive equilibrium point of $(\widetilde{\mathcal{G}}, \widetilde{\kappa})$, and therefore complex balanced by Theorem 3.3. The proof is concluded by (3.2) and by noting that, for any complex $\widetilde{y} \in \mathcal{C}_y$,

$$\{\widetilde{y}' \in \widetilde{\mathcal{C}} \colon \widetilde{y} \to \widetilde{y}' \in \widetilde{\mathcal{R}}\} = \{\widetilde{y}' \in \mathcal{C}_y \colon \widetilde{y} \to \widetilde{y}' \in \mathcal{R}_y\},$$
$$\{\widetilde{y}' \in \widetilde{\mathcal{C}} \colon \widetilde{y}' \to \widetilde{y} \in \widetilde{\mathcal{R}}\} = \{\widetilde{y}' \in \mathcal{C}_y \colon \widetilde{y}' \to \widetilde{y} \in \mathcal{R}_y\}.$$

□

B.2 Proof of Proposition 4.2

If $\mathcal{R}_\Gamma$ is empty there is nothing to prove. Suppose that this is not the case. Since $\mathcal{G}_\Gamma$ has deficiency zero, by Lemma A.2, it is weakly reversible. For any $y \to y' \in \mathcal{R}_\Gamma$, by definition there exists $x \in \Gamma$ such that $x \geq y$, which in turn implies $x + y' - y \geq y'$. Therefore, for any directed path in the reaction graph of $\mathcal{G}$ that starts with $y \to y' \in \mathcal{R}_\Gamma$, all the reactions in the path belong to $\mathcal{R}_\Gamma$, by definition of $\mathcal{R}_\Gamma$. Since $\mathcal{G}_\Gamma$ is weakly reversible, this can only happen if $\mathcal{R}_\Gamma \subseteq \mathcal{R}^*$, and this proves the first part of the statement. To conclude the proof, note that if the deficiency of $\mathcal{G}$ is zero, then by Lemma 3.2 the deficiency of $\mathcal{G}_\Gamma$ is zero as well. □

B.3 Proof of Theorem 4.3

For the first part of the statement, consider a continuous-time Markov chain $C_\Gamma(t)$ with state space $\Gamma \times \mathcal{C}$ and transition rate from (x, y) to $(x + y' - y, y')$ given by $\lambda_{y \to y'}(x)$ if $y \to y' \in \mathcal{R}_\Gamma$, and zero otherwise. The master equation for $C_\Gamma(t)$ is

$$\sum_{y \in \mathcal{C}_\Gamma} \tilde{\pi}(x - y' + y, y) \lambda_{y \to y'}(x - y' + y) = \sum_{y \in \mathcal{C}_\Gamma} \tilde{\pi}(x, y') \lambda_{y' \to y}(x) \quad \forall y' \in \mathcal{C}, x \in \Gamma,$$

with the convention that $\lambda_{y\to y'}(x) = 0$ if $y \to y' \notin \mathcal{R}_\Gamma$. By Definition 8, a stationary distribution for $C_\Gamma(t)$ exists and it is of the form $\tilde{\pi}(x, y) = M\pi(x)$, for a suitable normalising constant M. Since $\pi(x)$ is positive for any $x \in \Gamma$ (because it is a stationary distribution on an irreducible component), then by standard Markov chain theory, we have that for any two states $(x_1, y_1), (x_2, y_2) \in \Gamma \times \mathcal{C}$, if (x_2, y_2) is accessible from (x_1, y_1), then (x_1, y_1) is accessible from (x_2, y_2). Fix $y \to y' \in \mathcal{R}_\Gamma$ and $x \in \Gamma$ with $x \geq y$. Then, a directed path from $(x+y'-y, y')$ to (x, y) exists in the graph associated with $C_\Gamma(t)$. The second components of the form y of the states in the path, by construction, determine a directed path in the reaction graph of $\mathcal{G}_\Gamma$ from y' to y. Hence, any reaction $y \to y' \in \mathcal{R}_\Gamma$ is contained in a closed directed path, which means that $\mathcal{G}_\Gamma$ is weakly reversible.

Assume now that K is mass-action kinetics with rate constants κ and that c is a complex balanced equilibrium of $(\mathcal{G}, \kappa)$. Then, by Theorem 4.1, there exists a (unique) stationary distribution on Γ of the form (4.3). If a species S_j is not in $\mathcal{X}_\Gamma$, then the value of x_j is constant for any $x \in \Gamma$, and (4.5) can be obtained from (4.3) by modifying the normalising constant.

By Theorem 3.1 and Lemma A.3, we have that

$$\sum_{y\in\mathcal{C}_\Gamma} c^{y-y'}\kappa_{y\to y'} = \sum_{y\in\mathcal{C}_\Gamma} \kappa_{y'\to y} \quad \forall y' \in \mathcal{C}_\Gamma,$$

with $\kappa_{y\to y'} = 0$ if $y \to y' \notin \mathcal{R}_\Gamma$. Therefore, for any $y' \in \mathcal{C}_\Gamma$ and $x \in \Gamma$,

$$\frac{1}{(x-y')!}\sum_{y\in\mathcal{C}_\Gamma} c^{x+y-y'}\kappa_{y\to y'}\mathbb{1}_{\{x\geq y'\}} = \frac{1}{(x-y')!}\sum_{y\in\mathcal{C}_\Gamma} c^{x}\kappa_{y'\to y}\mathbb{1}_{\{x\geq y'\}},$$

which leads to (4.4), since π is of the form (4.3).

To prove the converse we first introduce a new stochastic mass-action system $(\hat{\mathcal{G}}_\Gamma, \hat{\kappa}_\Gamma)$, which is given by the reactions of the form

$$y + S_y \to y' + S_{y'} \quad \text{with } y \to y' \in \mathcal{R}_\Gamma,$$

where S_y are fictitious species in one to one correspondence with the complexes $\mathcal{C}_\Gamma$. The rate constant of the reaction $y + S_y \to y' + S_{y'}$ is given by $\kappa_{y\to y'}$. It is not difficult to see that the sum of the fictitious species is conserved for any possible trajectory. Moreover, since any directed path $y_1 \to y_2 \to \dots y_q$ in the reaction graph of $\mathcal{G}$ corresponds to a directed path $y_1 + S_{y_1} \to y_2 + S_{y_2} \to \dots y_q + S_{y_q}$ in the reaction graph of $\hat{\mathcal{G}}_\Gamma$, we have that $\hat{\mathcal{G}}_\Gamma$ is weakly reversible by the first part of the proof.

Consider the set

$$\Upsilon = \{(x, \hat{x}) \in \mathbb{N}^n \times \mathbb{N}^m : x \in \Gamma, \|\hat{x}\|_1 = 1\}.$$

Every state in Υ is of the form $(x, S_y) \in \mathbb{N}^{n+m}$, where $x \in \Gamma$ and S_y is considered as the vector in $\mathbb{N}^m$ with entry 1 in the position corresponding to the species S_y and 0 otherwise. Since Γ is an irreducible component of $\mathcal{G}$ and the sum of the fictitious species is conserved, no state outside Υ is accessible from any state in Υ, according to $\hat{\mathcal{G}}_\Gamma$. Moreover, the master equation on Υ can be written as

$$\sum_{y\in\mathcal{C}_\Gamma} \hat{\pi}(x-y'+y, S_y)\kappa_{y\to y'}\frac{(x-y'+y)!}{(x-y')!}\mathbb{1}_{\{x\geq y'\}} = \sum_{y\in\mathcal{C}_\Gamma} \hat{\pi}(x, S_{y'})\kappa_{y'\to y}\frac{x!}{(x-y')!}\mathbb{1}_{\{x\geq y'\}} \quad \forall y' \in \mathcal{C}, x \in \Gamma. \quad \text{(B.1)}$$

If we choose $\hat{\pi}(x, \hat{x}) = M\pi(x)$ for some positive constant M, then the master equation (B.1) is satisfied due to Definition 8. Therefore, if M is chosen as a suitable normalising constant, $\hat{\pi}(x, z) = M\pi(x)$ is a stationary distribution on Υ.

Consider the linear homomorphism φ as defined in (3.3), for the reaction network $\hat{\mathcal{G}}_\Gamma$. Note that $|\hat{\mathcal{C}}_\Gamma| = |\mathcal{C}_\Gamma| = m_\Gamma$. For any vector e_y of the basis of $\mathbb{R}^{m_\Gamma}$, we have $\varphi(e_y) = (y, S_y)$. Since the vectors (y, S_y) with $y \in \mathcal{C}_\Gamma$ are linear independent, φ is an isomorphism and the deficiency of $\hat{\mathcal{G}}_\Gamma$ is 0.

Since $\hat{\mathcal{G}}_\Gamma$ is a deficiency zero weakly reversible reaction network, it follows from Theorem 3.3 that the mass-action system $(\hat{\mathcal{G}}_\Gamma, \kappa)$ is complex balanced. Therefore, by Theorem 4.1, we have that $\hat{\pi}$ has the form

$$\hat{\pi}(x, \hat{x}) = M_{\hat{\Gamma}}^{(c,\hat{c})}\frac{c^x}{x!}\frac{\hat{c}^{\hat{x}}}{\hat{x}!},$$

for a complex balanced equilibrium $(c, \hat{c})$, on any irreducible component $\hat{\Gamma}$ contained in Υ. Since $\hat{\pi}(x, \hat{x}) = M\pi(x)$ does not depend on $\hat{x}$, we have

$$\hat{\pi}(x, \hat{x}) = M_\Gamma^c \frac{c^x}{x!},$$

for any $(x, \hat{x}) \in \Upsilon$.

Fix a complex $y' \in \mathcal{C}_\Gamma$. Since $\mathcal{G}_\Gamma$ is weakly reversible, there exists a reaction $y' \to y$ that is active on Γ. Fix $x \in \Gamma$ such that $x \geq y'$. Then for any $y \to y' \in \mathcal{R}_\Gamma$ we have $x - y' + y \geq y$. If we plug the formula for $\hat{\pi}(x, \hat{x})$ in (B.1) for our choice of x and y', we obtain

$$\sum_{y \in \mathcal{C}_\Gamma} M_\Gamma^c \frac{c^{x-y'+y}}{(x-y'+y)!} \kappa_{y \to y'} \frac{(x-y'+y)!}{(x-y')!} = \sum_{y \in \mathcal{C}_\Gamma} M_\Gamma^c \frac{c^x}{x!} \kappa_{y' \to y} \frac{x!}{(x-y')!},$$

which leads to

$$\sum_{y \in \mathcal{C}_\Gamma} c^{y-y'} \kappa_{y \to y'} = \sum_{y \in \mathcal{C}_\Gamma} \kappa_{y' \to y}.$$

The proof is concluded by the fact that the above holds for any fixed $y' \in \mathcal{C}_\Gamma$, which means that c is a complex balanced equilibrium of $(\mathcal{G}_\Gamma, \kappa_\Gamma)$. □

B.4 Proof of Theorem 4.6

By Lemma A.2, $\mathcal{G}_\Gamma$ is weakly reversible. Moreover, for $y \to y' \in \mathcal{R}_\Gamma$, if $x \geq y$ then $x + y' - y \geq y'$. This implies that for any directed path in the reaction graph of $\mathcal{G}$ that starts with $y \to y' \in \mathcal{R}_\Gamma$, all the reactions in the path belong to $\mathcal{R}_\Gamma$, by definition of $\mathcal{R}_\Gamma$. Since $\mathcal{G}_\Gamma$ is weakly reversible, every directed path in the reaction graph of $\mathcal{G}$ that starts with $y \to y' \in \mathcal{R}_\Gamma$ is contained in a closed directed path. This implies that $\mathcal{R}_\Gamma \subseteq \mathcal{R}^*$, and proves the first part of the statement.

Now assume that K is mass-action kinetics with rate constants κ. If the deficiency of $\mathcal{G}$ is zero, then by Lemma 3.2 the deficiency of the terminal network is zero as well. Moreover, $\mathcal{G}^*$ is weakly reversible by definition, thus by Theorem 3.3 $(\mathcal{G}^*, \kappa^*)$ is complex balanced for any choice of rate constants κ^*.

Let $X(t)$ be the stochastic process associated with $(\mathcal{G}, \kappa)$. By the first part of the statement, on Γ only terminal reactions take place and these involve a subset of the species only. Without loss of generality, we can assume that $\mathcal{X}^*$ is constituted by the first n^* species of $\mathcal{X}$. Therefore, Γ is of the form $\Gamma^* \times \{v\}$, with $\Gamma^* \subseteq \mathcal{R}^{n^*}$ and $v \in \mathbb{R}^{n-n^*}$. Moreover, we have that on Γ^*, the projection $X^*(t) = (X_1(t), \ldots, X_{n^*}(t))$ is distributed as the process associated with $(\mathcal{G}^*, \kappa^*)$, for which Γ^* is an irreducible component. Let c be a complex balanced equilibrium for $(\mathcal{G}^*, \kappa^*)$. Hence, by Theorem 4.1 or Corollary 4.4, the stationary distribution of the process $X(t) = (X^*(t), v)$ on Γ is of the form (4.6). □

B.5 Proof of Theorem 4.7

For the first part, we prove that if an irreducible component Γ is positive, then $\mathcal{G}$ is weakly reversible. This simply follows from Lemma A.2: indeed, by the lemma, $\mathcal{G}_\Gamma$ is weakly reversible and since Γ is positive, $\mathcal{G}_\Gamma = \mathcal{G}$.

To prove the second part, we have to show that a weakly reversible reaction network is essential, and this is done in [21]. Moreover, a deficiency zero weakly reversible mass-action system is complex balanced, and the proof is concluded by Theorem 4.1 or Corollary 4.4. □

References

[1] David F Anderson, Gheorghe Craciun, Manoj Gopalkrishnan, and Carsten Wiuf, *Lyapunov functions, stationary distributions, and non-equilibrium potential for chemical reaction networks.* arXiv:1410.4820, 2014.

[2] David F Anderson, Gheorghe Craciun, and Thomas G Kurtz, *Product-form stationary distributions for deficiency zero chemical reaction networks*, Bull. Math. Biol., 72 (2010), pp. 1947–1970.

[3] David F Anderson, Germán A Enciso, and Matthew D Johnston, *Stochastic analysis of biochemical reaction networks with absolute concentration robustness*, J. of the Royal Society Interface, 11 (2014), p. 20130943.

[4] David F Anderson and Thomas G Kurtz, *Continuous time Markov chain models for chemical reaction networks*, in Design and Analysis of Biomolecular Circuits: Engineering Approaches to Systems and Synthetic Biology, Heinz Koeppl, Douglas Densmore, Gianluca Setti, and Mario Di Bernardo, eds., Springer, 2011, pp. 3–42.

[5] Karen Ball, Thomas G Kurtz, Lea Popovic, and Greg Rempala, *Asymptotic analysis of multiscale approximations to reaction networks*, Ann. Appl. Probab., 16 (2006), pp. 1925–1961.

[6] L Boltzmann, *Neuer beweis zweier sätze über das wärmegleichgewicht unter mehratomigen gasmolekülen*, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien, 95 (1887), pp. 153–164.

[7] Gheorghe Craciun and Casian Pantea, *Identifiability of chemical reaction networks*, J. Math. Chem., 44 (2008), pp. 244–259.

[8] Péter Érdi and János Tóth, *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*, Manchester University Press, 1989.

[9] Martin Feinberg, *Chemical reaction network structure and the stability of complex isothermal reactors-I. The deficiency zero and deficiency one theorems*, Chemical Engineering Science, 42 (1987), pp. 2229–2268.

[10] Ankit Gupta and Mustafa Khammash, *Determining the long-term behavior of cell populations: A new procedure for detecting ergodicity in large stochastic reaction networks*, in Proceedings of the 19th IFAC World Congress, 2014.

[11] D Gusak, A Kukush, A Kulik, Y Mishura, and A Pilipenko, *Theory of Stochastic Processes: With Applications to Financial Mathematics and Risk Theory*, Springer New York, 2010.

[12] K L Hey, H Momiji, K Featherstone, J R E Davis, M R H White, and D. A. Rand, *A stochastic transcriptional switch model for single cell imaging data*, Biostatistics, in press (2015).

[13] Fritz Horn and Roy Jackson, *General mass action kinetics*, Arch. Ration. Mech. Anal., 47 (1972), pp. 81–116.

[14] Piers J Ingram, Michael PH Stumpf, and Jaroslav Stark, *Nonidentifiability of the source of intrinsic noise in gene expression from single-burst data*, PLoS Computational Biology, 4 (2008), p. e1000192.

[15] J R Jackson, *Networks of waiting lines*, Oper. Res., 5 (1957), pp. 518–521.

[16] Hye-Won Kang and Thomas G Kurtz, *Separation of time-scales and model reduction for stochastic reaction networks*, Ann. Appl. Probab., 23 (2013), pp. 529–583.

[17] Frank P Kelly, *Reversibility and stochastic networks*, John Wiley & Sons, Inc., 1979.

[18] Thomas G Kurtz, *The relationship between stochastic and deterministic models for chemical reactions*, J. Chem. Phys., 57 (1972), pp. 2976–2978.

[19] Jean Mairesse and Hoang-Thach Nguyen, *Deficiency zero petri nets and product form*, in Applications and Theory of Petri Nets, Springer, 2009, pp. 103–122.

[20] Andrea Marin, Simonetta Balsamo, and Peter G Harrison, *Analysis of stochastic Petri nets with signals*, Performance Evaluation, 69 (2012), pp. 551–572.

[21] Loïc Paulevé, Gheorghe Craciun, and Heinz Koeppl, *Dynamical properties of discrete reaction networks*, J. Math. Biol., 69 (2014), pp. 55–72.

[22] Peter Whittle, *Systems in stochastic equilibrium*, John Wiley & Sons, Inc., 1986.

[23] C Zechner, M Unger, S Pelet, M Peter, and H Koeppl, *Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings*, Nat. Methods, 11 (2014), pp. 197–202.

III

Finite time behavior of stochastically modeled chemical systems with absolute concentration robustness

David F. Anderson
Department of Mathematics
University of Wisconsin-Madison

Daniele Cappelletti
Department of Mathematical Sciences
University of Copenhagen

Thomas G. Kurtz
Department of Mathematics
University of Wisconsin-Madison

Publication details

Unsubmitted.

Finite time behavior of stochastically modeled chemical systems with absolute concentration robustness

David F. Anderson* Daniele Cappelletti† Thomas G. Kurtz‡

September 17, 2015

Abstract

Recent work pertaining to both deterministic and stochastic models of chemical reaction systems satisfying an "absolute concentration robustness" (ACR) property has focused on the limiting behaviours of the models as time goes to infinity. In particular, it is known that in the stochastic setting an ACR system will undergo an extinction event with a probability of one so long as the system is conservative. Here we consider a general class of stochastic chemical models that intersects with the class of ACR systems. We consider a specific system scaling over compact time intervals and prove that in a limit of this scaling the distribution of the time-averaged abundances of the ACR species converges weakly to a certain time-average of an identified product-form Poisson distribution. This result is in agreement with recent conjectures pertaining to the behaviour of stochastically modeled ACR systems on compact time intervals.

1 Introduction

Biochemical reaction networks are often quite complex and computationally intractable. It is therefore important to develop mathematical techniques that relate simple graphical features of the reaction network, which are easy to check, to the qualitative dynamics of the underlying mathematical model. This approach dates back to at least [9, 10, 11], where graphical characteristics of the networks ensure uniqueness and local asymptotic stability of the steady states for deterministically modeled complex-balanced systems.

In this context, Shinar and Feinberg provided graphical conditions that imply certain species satisfy an *absolute concentration robustness* (ACR) property for the associated deterministically modeled system [20]. A species is said to posses ACR if for a fixed choice of system parameters its concentration is the same at any positive equilibrium point of the deterministically modeled system. Such a feature has been observed in several important biochemical reaction networks, including signal transduction cascades and gene regulatory networks [6, 20]. The ACR property provides useful information on the system dynamics since it indicates a predictable fixed response regardless of changes in the environment. Followup research pertaining to deterministically modeled systems with ACR species can be found in [14].

Stochastically modeled systems satisfying essentially the same graphical conditions as those detailed in [20] were considered by Anderson, Enciso, and Johnston in [3]. There it was shown that stochastically modeled ACR systems undergo an extinction event with a probability of one, so long as the system is conservative. Moreover, in the examples shown in [3] the total mass is acquired by the species possessing ACR. Such a result can be considered an example of a discrepancy between the limiting behaviour of a deterministic system and the limiting behaviour of the corresponding stochastic system, with one modeling choice predicting a form of long-term stability and the other predicting long-term instability. However, in [3] it is pointed out that the extinction event is typically a rare event on reasonable timeframes and

*anderson@math.wisc.edu
†Corresponding author. d.cappelletti@math.ku.dk
‡kurtz@math.wisc.edu

that useful information pertaining to the behaviour of stochastically modeled ACR systems could be had by better understanding the dynamics of the system on compact time intervals or via the quasi-stationary distribution. It is conjectured in [3] that the distribution of the ACR species will be approximately Poisson in either case. Both a simple example pertaining to a model of protein interactions and a numerical analysis of the two-component EnvZ/OmpR signaling system in Escherichia coli provide evidence in favor of the conjecture [3].

In this paper we provide an asymptotic result for a class of stochastically modeled systems that overlaps with ACR reaction systems. In particular, we consider a multiscale setting in which the abundances of the ACR species are of order $O(1)$, while the abundances of the other species are all of order $O(N)$. We then scale the rate constants in a particular way and let N going to infinity. Under this limit, we prove that on compact time intervals the ACR species behave in the way conjectured in [3]. Namely, the distribution of the ACR species is approximated on average by a product-form Poisson distribution whose parameter is given by the ACR equilibrium value of an associated deterministically modeled system. Thus, the results presented here link the qualitative behaviours of the deterministic and stochastic models. Furthermore, the result fully explains the outcome of the numerical analysis of the EnvZ/OmpR signaling system performed in [3].

We end this section with two instructive examples that demonstrate our main results.

Example 1.1. Consider the deterministically modeled system with reaction network

$$\begin{aligned} A + B &\xrightarrow{\kappa_1} 2B \\ B &\xrightarrow{\kappa_2} A \end{aligned} \tag{1}$$

and mass action kinetics (see (6) and (7)). The species A exhibits ACR since the equilibrium amount of A is always κ_2/κ_1, regardless of the concentration of species B [3, 20].

Now consider a sequence $\{X^N\}_{N\in\mathbb{N}}$ of continuous time Markov chain models for (1), in which the counts of species A and B at time t are given by $X_1^N(t)$ and $X_2^N(t)$, respectively. We suppose the Nth model has initial condition

$$(X_1^N(0), X_2^N(0)) = (a_N, N(b + \epsilon_N)) \in \mathbb{Z}^2_{\geq 0},$$

where $(a_N)_{N\in\mathbb{N}}$ is a bounded sequence of natural numbers, b is a positive real number and $(\varepsilon_N)_{N\in\mathbb{N}}$ is a sequence of real numbers tending to zero. This choice corresponds to an experiment where the abundance of the molecules of B is increased, while the magnitude of the count of A is maintened. Our goal will be to understand the limiting behaviour of X_1^N for N going to infinity.

Let J be a random variable having a Poisson distribution with parameter κ_2/κ_1. Corollary 4.1 will allow us to conclude that for any continuous function $g : \mathbb{N} \to \mathbb{R}$ with at most polynomial growth rate, the distribution of $g(X_1^N(t))$ converges on average to the distribution of $g(J)$ for N going to infinity. Specifically, for any real positive T, we have

$$\int_0^{\cdot\wedge T} \Big(g(X_1^N(s)) - E[g(J)]\Big) ds \xrightarrow[N\to\infty]{} 0.$$

We refer to Example 4.1 for more details. □

Example 1.2. Our results are also applicable to models that do not utilize mass action kinetics. Consider the stochastic reaction system

$$\begin{aligned} A + 2B &\longrightarrow 3B \\ B &\underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} C \xrightarrow{\kappa_3} A, \end{aligned} \tag{2}$$

where the rate of the reaction $A + 2B \to 3B$ is given by

$$\lambda(x) = \kappa_0 \frac{x_1 x_2 (x_2 - 1)}{1 + x_2}, \tag{3}$$

and where $\kappa_0 \in \mathbb{R}_{>0}$. The rate (3) corresponds to an inhibitory effect of the molecules of B on the production of B itself. If we consider a sequence of such models in which the counts of B and C go to infinity, then the limiting behaviour of the model (2)-(3) coincides with the limiting behaviour of the process associated with the reaction network

$$\begin{aligned} A + B &\xrightarrow{\kappa_0} 2B \\ B &\underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} C \xrightarrow{\kappa_3} A. \end{aligned} \tag{4}$$

Due to [20], the reaction system (4) exhibits ACR in the species A, when deterministically modeled. Let $q \in \mathbb{R}_{>0}$ be the ACR value for species A. Due to the connection between the models (2)-(3) and (4), we anticipate the value q will play a role in the limiting behaviour of species A of (2)-(3).

We therefore denote by $\{X^N\}_{N\in\mathbb{N}}$ a sequence of stochastic processes modeled according to (2)-(3), with $X_1^N(t)$, $X_2^N(t)$, and $X_3^N(t)$ being the counts at time t for the species A, B, and C, respectively. We suppose the Nth model satisfies

$$X^N(0) = (a_N, N(b+\varepsilon_{1,N}), N(c+\varepsilon_{2,N})) \in \mathbb{Z}^3_{\geq 0},$$

where $(a_N)_{N\in\mathbb{N}}$ is a bounded sequence of natural numbers, b, c are positive real numbers, and $(\varepsilon_{1,N})_{N\in\mathbb{N}}$ and $(\varepsilon_{2,N})_{N\in\mathbb{N}}$ are sequences of real numbers tending to zero. Our aim is to understand the limiting behaviour of X_1^N as N goes to infinity.

Let J be a Poisson distribution with parameter q, the ACR value of for species A when (4) is modeled deterministically. By Corollary 4.2, we have that for any continuous function g with at most polynomial growth rate, after a certain time and for N large enough, the distribution of $g(X_1^N(t))$ can be approximated on average by the distribution of $g(J)$. Specifically, for any $\varepsilon > 0$ there exists $H_\varepsilon > 0$ such that, for any $t_2 > t_1 > H_\varepsilon$

$$\lim_{N\to\infty} P\left(\sup_{t\in[t_1,t_2]} \left| \int_{t_1}^{t} \Big(g(X_1^N(s)) - E[g(J)] \Big) ds \right| > \varepsilon(1 + t_2 - t_1) \right) = 0.$$

See Example 4.2 for more details. □

2 Necessary Background and Notation

For any real vector v, we denote its ith entry by v_i. We will write $v > 0$ if every entry of v is strictly positive. We denote by $[v]$ the vector of the floor functions of the entries of v; that is, $[v]_i = \lfloor v_i \rfloor$. For any real vector α of the same size as v, and for $N > 0$, we denote by $N^\alpha v$ the vector satisfying

$$(N^\alpha v)_i = N^{\alpha_i} v_i.$$

We will denote by $\|v\|$ the euclidean norm of the vector, by $\|v\|_1$ its L^1-*norm* and by $\|v\|_\infty$ its L^∞-*norm*, that is

$$\|v\| = \sqrt{\sum_i v_i^2}, \qquad \|v\|_1 = \sum_i |v_i|, \quad \text{and} \quad \|v\|_\infty = \max_i |v_i|.$$

For two vectors v and w of the same dimension, we write $v < w$, $v \leq w$, $v > w$ or $v \geq w$ if the inequality holds component-wise. Furthermore, for any set A we will indicate by $|A|$ its cardinality and by $\mathbb{1}_A$ its indicator function. Finally, for any $a, b \in \mathbb{R}$, $a \wedge b$ and $a \vee b$ will denote $\min\{a, b\}$ and $\max\{a, b\}$, respectively.

We say that a function $g\colon \mathbb{R}^n \to \mathbb{R}$ has at most *polynomial growth rate* if there exists a multivariate polynomial $p\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\limsup_{\|x\|\to\infty} \frac{g(x)}{p(x)} = 0$$

Here we give some basic definitions from chemical reaction network theory, see for example [7, 9] for a more detailed introduction.

A *reaction network* is a triple $\mathcal{G} = (\mathcal{X}, \mathcal{C}, \mathcal{R})$. $\mathcal{X}$ is a finite non-empty ordered set of symbols, referred to as *species*, and $\mathcal{C}$ is a finite non-empty ordered set of linear combinations of species on $\mathbb{N}$, referred to as *complexes*. Any species $S_i \in \mathcal{X}$ can be identified with the vector $e_i \in \mathbb{R}^{|\mathcal{X}|}$, whose ith entry is 1 and whose other entries are zero. Therefore, any complex $y \in \mathcal{C}$ will be identified with a vector in $\mathbb{R}^{|\mathcal{X}|}$ that is linear combination of the vectors e_i. Finally, $\mathcal{R}$ is a non-empty ordered subset of $\mathcal{C} \times \mathcal{C}$, whose elements are called *reactions*, such that for any $y \in \mathcal{C}$, $(y, y) \notin \mathcal{R}$. Following the common notation, we will denote any element $(y_r, y_r') \in \mathcal{R}$ by $y_r \to y_r' \in \mathcal{R}$, in which case we then call y_r the *source complex* and y_r' the *product complex* of that reaction. It is possible that a complex $y \in \mathcal{C}$ is the source (product) complex of different reactions, and that it is both the source complex of one reaction and the product complex of another reaction. It is commonly required that every species $S \in \mathcal{X}$ appears in at least one complex, and that every complex $y \in \mathcal{C}$ appears as an element in at least one reaction. It is possible to associate a directed graph to $\mathcal{G}$, where the set of nodes is the set of complexes $\mathcal{C}$ and the arrows are given by the reactions $y_r \to y_r' \in \mathcal{R}$. If the graph is such that for any directed path from y to y' there exists a directed path from y' to y, then $\mathcal{G}$ is *weakly reversible*. For the rth reaction, $y_r \to y_r'$, we denote by $\xi_r = y_r' - y_r$ the corresponding *reaction vector*. We write $S_i \in \xi_r$ ($S_i \in y_r$) if $\xi_{ri} \neq 0$ ($y_{ri} \neq 0$). For any species $S \in \mathcal{X}$, let

$$\mathcal{R}_S := \{y_r \to y_r' \in \mathcal{R} \,:\, S \in \xi_r\}, \tag{5}$$

the set of reactions that change the amount of species S.

To each reaction $y_r \to y_r' \in \mathcal{R}$, we can associate a function $\lambda_r : \mathbb{R}^{|\mathcal{X}|}_{\geq 0} \to \mathbb{R}_{\geq 0}$. The set consisting of these function $\mathcal{K} = \{\lambda_r\}_{y_r \to y_r' \in \mathcal{R}}$ is referred to as the *kinetics*, and the functions λ_r are called *rate functions*, or *intensity functions*, or *propensity functions*. The pair $\mathscr{S} = (\mathcal{G}, \mathcal{K})$ is a *reaction system*, which can be stochastically or deterministically modeled, as explained below.

In a *stochastically modeled reaction system* $\mathscr{S} = (\mathcal{G}, \mathcal{K})$, the counts of molecules of the different chemical species are considered, and the counts at time t form a vector $X(t) \in \mathbb{N}^{|\mathcal{X}|}$. The evolution in time of the vector $X(t)$ follows a continuous time Markov chain, where in each state $x \in \mathbb{N}^{|\mathcal{X}|}$ the obtainable states are $\{x + \xi_r\}$ and transition rates are given by $\{\lambda_r(x)\}$, with $y_r \to y_r'$ varying in $\mathcal{R}$. If at time t^* the reaction $y_r \to y_r'$ occurs, then we have

$$X(t^*) = X(t^*-) + \xi_r,$$

where $X(t^*-)$ denotes the previous state. To stick with the physical meaning of the reactions, we require that the kinetics is such that for any reaction $y_r \to y_r' \in \mathcal{R}$ we have $\lambda_{y_r \to y_r'}(x) > 0$ only if $x \geq y_r$. This condition prevents the number of molecules present from becoming negative. Moreover, in this setting, we are only interested in the values of $\lambda_{y_r \to y_r'}(x)$, when $x \in \mathbb{N}^{|\mathcal{X}|}$, therefore the domain of the rate function can be restricted to $\mathbb{N}^{|\mathcal{X}|}$. Following the terminology utilized in [4, 8, 15], we can write

$$X(t) = X(0) + \sum_{y_r \to y_r' \in \mathcal{R}} Y_r \left(\int_0^t \lambda_r(X(s)) ds \right) \xi_r,$$

where the Y_r are i.i.d. unit rate Poisson processes. For any two states $x, z \in \mathbb{N}^{|\mathcal{X}|}$, we say that a state z is *obtainable* from x if there exists a sequence of reactions $(y_{r_i} \to y_{r_i}')_{i=1}^m$ such that

$$z = x + \sum_{i=1}^m \xi_{r_i},$$

and for which

$$\lambda_{y \to y'} \left(x + \sum_{j=1}^i \xi_{r_j} \right) > 0,$$

for all $i \in \{0, \ldots, m-1\}$. We further say that $\mathcal{G}$ is *irreducible* if for any two states $x, z \in \mathbb{N}^{|\mathcal{X}|}$, z is obtainable from x and x is obtainable from z. See [18] for more on irreducible reaction networks and for

sufficient conditions implying irreducibility. A popular choice of kinetics for stochastic reaction systems is given by *stochastic mass action kinetics*, defined by

$$\lambda_r(x) = \kappa_r \frac{x!}{(x-y_r)!} \mathbb{1}_{\{x \geq y_r\}},$$

where $\kappa_r \in \mathbb{R}_{>0}$ are called *rate constants* and for any vector $v \in \mathbb{N}^m$, $v!$ is defined by

$$v! = \prod_{i=1}^{m} v_i!,$$

with the convention $0! = 1$. This kinetics is related to the assumption that the system is well-stirred, so the propensity of each reaction is proportional to the possible sets of present molecules that can give rise to an occurrence of the reaction. A stochastic reaction system endowed with stochastic mass action kinetics is referred to as *stochastic mass action system*, and will be denoted $\mathscr{S} = (\mathcal{G}, \kappa)$.

In a *deterministically modeled reaction system* $\mathscr{S} = (\mathcal{G}, \mathcal{K})$, the concentrations of the different chemical species are considered, and the concentrations at time t form a vector $z(t) \in \mathbb{R}^{|\mathcal{X}|}_{\geq 0}$. The evolution in time of the vector $z(t)$ obeys the ODE

$$z'(t) = \sum_{y_r \to y_r' \in \mathcal{R}} \xi_r \lambda_r(z(t)), \tag{6}$$

or in integral form

$$z(t) = z(0) + \sum_{y_r \to y_r' \in \mathcal{R}} \xi_r \int_0^t \lambda_r(z(s)) ds.$$

As in the stochastic case, we put some restriction on the kinetics and require that for any $y_r \to y_r' \in \mathcal{R}$ we have $\lambda_{y_r \to y_r'}(x) > 0$ only if $x_i > 0$ whenever $S_i \in y_r$. This condition means that a reaction cannot take place if some necessary chemical species is missing, and it guarantees that the vector $z(t)$ will remain non-negative. *Deterministic mass action kinetics* is given by

$$\lambda_r(x) = \kappa_r x^{y_r}, \tag{7}$$

where $\kappa_r \in \mathbb{R}_{>0}$ are called *rate constants* and for any two vectors $v, w \in \mathbb{N}^m$, v^w is defined by

$$v^w = \prod_{i=1}^{m} v_i^{w_i},$$

with the convention $0^0 = 1$. Thus, the rate of each reaction is proportional to the products of the concentrations of the species appearing in the source complex, according to multiplicity. As in the stochastic case, this kinetics is chosen for well-stirred systems. A deterministic reaction system with deterministic mass action kinetics is termed a *deterministic mass action system*, and will be denoted by $\mathscr{S} = (\mathcal{G}, \kappa)$.

A fruitful notion in chemical reaction network theory, and one that will play a role in the present work, is that of a *complex balanced equilibrium*, which is a positive equilibrium point c of a deterministic mass action system satisfying

$$\sum_{\substack{y_r \to y_r' \in \mathcal{R} \\ y_r = y}} \kappa_r c^{y_r} = \sum_{\substack{y_r \to y_r' \in \mathcal{R} \\ y_r' = y}} \kappa_r c^{y_r} \quad \text{for each } y \in \mathcal{C},$$

where the sum on the left, respectively right, is over those reactions for which y is the source, respectively product, complex. We say that a deterministic mass action system is *complex balanced* if there exists at least one positive equilibrium point, and if every positive equilibrium point is a complex balanced equilibrium.

We extend the definition of complex balanced to the stochastic setting by saying that a stochastic mass action system $(\mathcal{G}, \kappa)$ is complex balanced if the deterministic mass action system $(\mathcal{G}, \kappa)$ is complex balanced. We may therefore refer to complex balanced mass action systems without specifying whether they are

stochastically or deterministically modeled. In the same fashion, whenever we refer to an equilibrium point of a reaction system, we implicitly assume it is an equilibrium point for the deterministically modeled system.

It is worth noting that under the assumptions detailed above for both deterministic and stochastic reaction systems, the evolution of the amounts of species present is restricted to

$$X(t) \in \left(X(0) + \operatorname{span}\{\xi_r\}_{y_r \to y_r' \in \mathcal{R}}\right) \cap \mathbb{Z}_{\geq 0}^{|\mathcal{X}|} \quad \text{and} \quad z(t) \in \left(z(0) + \operatorname{span}\{\xi_r\}_{y_r \to y_r' \in \mathcal{R}}\right) \cap \mathbb{R}_{\geq 0}^{|\mathcal{X}|},$$

regardless of the choice of kinetics $\mathcal{K}$. The sets $\left(v + \operatorname{span}\{\xi_r\}_{y_r \to y_r' \in \mathcal{R}}\right)$ with $v \in \mathbb{R}^{|\mathcal{X}|}$ are called the *stoichiometric compatibility classes* of $\mathcal{G}$, and the sets $\left(v + \operatorname{span}\{\xi_r\}_{y_r \to y_r' \in \mathcal{R}}\right) \cap \mathbb{R}_{\geq 0}^{|\mathcal{X}|}$ and $\left(v + \operatorname{span}\{\xi_r\}_{y_r \to y_r' \in \mathcal{R}}\right) \cap \mathbb{R}_{>0}^{|\mathcal{X}|}$ are called the *non-negative stoichiometric compatibility classes* and *positive stoichiometric compatibility classes* of $\mathcal{G}$. Any vector $T \in \mathbb{R}^{|\mathcal{X}|}$ that is orthogonal to the stoichiometric compatibility classes of $\mathcal{G}$ is a *conservation law* for $\mathcal{G}$, and if there exists a positive conservation law for $\mathcal{G}$, then $\mathcal{G}$ is called *conservative*.

Let

$$s := \dim\left(\operatorname{span}\{\xi_r\}_{y_r \to y_r' \in \mathcal{R}}\right).$$

We define the *deficiency* of $\mathcal{G}$ as

$$\delta := |\mathcal{C}| - \ell - s,$$

where ℓ is the number of connected components of the directed graph associated with $\mathcal{G}$. We end this section by stating some classical results that can be found in [9, 10, 11], which connect graphical and dynamical features of the deterministic mass action systems and will be of use to us.

Theorem 2.1. *If a deterministic mass action system $\mathscr{S} = (\mathcal{G}, \kappa)$ possesses a complex balanced equilibrium, then $\mathscr{S}$ is complex balanced and $\mathcal{G}$ is weakly reversible. Moreover, there exists exactly one complex balanced equilibrium in every positive stoichiometric compatibility class, and it is locally asymptotically stable relative to its positive stoichiometric compatibility class.*

Theorem 2.2. *If $\mathcal{G}$ is weakly reversible and has deficiency 0, then for any choice of rate constants the deterministic mass action system $\mathscr{S} = (\mathcal{G}, \kappa)$ is complex balanced.*

3 The multiscale setting and main results

Denote by $\mathcal{K}^N$ a sequence of stochastic kinetics for $\mathcal{G}$, with $N \in \mathbb{N}_{>0}$, and let $X^N(t)$ be the sequence of stochastic processes associated with the system $(\mathcal{G}, \mathcal{K}^N)$. Assume that there exists a vector $\alpha \in \{0,1\}^{|\mathcal{X}|}$ such that

$$\lim_{N \to \infty} N^{-\alpha} X^N(0) = X_0 > 0. \tag{8}$$

The condition (8) implies a partition of the set of species $\mathcal{X}$ in two sets, the *discrete species* (denoted by $\mathcal{X}_d$) and the *continuous species* (denoted by $\mathcal{X}_c$),

- $S_i \in \mathcal{X}_d$ if $\alpha_i = 0$, in which case $X_i^N(0) = O(1)$;
- $S_i \in \mathcal{X}_c$ if $\alpha_i = 1$, in which case $X_i^N(0) = O(N)$.

Let

$$\pi_d \colon \mathbb{R}^{|\mathcal{X}|} \to \mathbb{R}^{|\mathcal{X}_d|} \quad \text{and} \quad \pi_c \colon \mathbb{R}^{|\mathcal{X}|} \to \mathbb{R}^{|\mathcal{X}_c|}$$

be the projections onto the discrete and continuous species, respectively, and define

$$X_{\text{disc}}^N(t) := \pi_d(X^N(t)) \quad \text{and} \quad X_{\text{cont}}^N(t) := \pi_c(X^N(t)).$$

For convenience, we will sometimes consider the rate functions as functions from $\mathbb{N}^{|\mathcal{X}_d|} \times \mathbb{N}^{|\mathcal{X}_c|}$, and write $\lambda_r(v, w)$, where v and w denote the amounts of the discrete and continuous species, respectively. For any reaction $y_r \to y_r' \in \mathcal{R}$, define

$$\beta_r = \max_{S_i \in y_r} \alpha_i.$$

We assume that

$$\lim_{N\to\infty} N^{-\beta_r}\lambda_r^N(v,[Nw]) = \lambda_r(v,w) \tag{9}$$

uniformly on the compact sets of $\mathbb{N}^{|\mathcal{X}_d|}\times\mathbb{R}_{\geq 0}^{|\mathcal{X}_c|}$, where the functions λ_r are non-zero and locally Lipschitz, with domain $\mathbb{N}^{|\mathcal{X}_d|}\times\mathbb{R}_{\geq 0}^{|\mathcal{X}_c|}$. We denote by $\mathcal{K}$ the kinetics given by the limiting functions λ_r. We define

$$\nu_r^N(v,w) := N^{-\beta_r}\lambda_r^N(v,[Nw]) - \lambda_r(v,w). \tag{10}$$

The above setting is a particular case of the one studied in [5, 12, 19].

Remark 3.1. If the kinetics $\mathcal{K}^N$ are stochastic mass action kinetics for all $N\in\mathbb{N}_{>0}$, then by (9) the sequence

$$N^{-\beta_r}\lambda_r^N(v,[Nw]) = N^{-\beta_r}\kappa_r^N\frac{v!}{(v-\pi_d(y_r))!}\frac{[Nw]!}{([Nw]-\pi_c(y_r))!}$$

converges to a positive number for some $(v,w)\in\mathbb{N}^{|\mathcal{X}_d|}\times\mathbb{R}_{\geq 0}^{|\mathcal{X}_c|}$. Since the sequence

$$N^{-\|\pi_c(y_r)\|_1}\frac{[Nw]!}{([Nw]-\pi_c(y_r))!}$$

converges to w^{y_r}, we have

$$\lim_{N\to\infty} N^{-\beta_r+\|\pi_c(y_r)\|_1}\kappa_r^N = \kappa_r, \tag{11}$$

for some positive constant κ_r.

Here we make some key assumptions.

Assumption 1. For any $S\in\mathcal{X}_d$, there exists at least one reaction $y_r\to y_r'\in\mathcal{R}_S$ such that $\beta_r = 1$ (i.e. the species S is *fast* consumed or produced, recall $\mathcal{R}_S$ defined in (5)).

Assumption 2. The discrete species only appear with stoichiometric coefficient 1.

In order to motivate and explain the above scaling and assumptions we consider sequences of processes satisfying the network structures of Examples 1.1 and 1.2.

Consider first the network of Example 1.1 with rate constants κ_1 and κ_2, and suppose that the total initial abundance of the system (i.e. the sum of the abundances of species A and B) is large, and that the system is near the known ACR equilibrium, in which case $X_1(0)\approx q = \kappa_2/\kappa_1$. Specifically, we suppose that $X_1^N(0)+X_2^N(0) = N$ for some large $N\in\mathbb{N}$ and that $X_1^N(0) = O(1)$ in N, in which case $X_2^N(0) = N - X_1^N(0) = O(N)$. We will be interested in letting $N\to\infty$. In this setting, A is a discrete species, $\alpha_1 = 0$, and B is a continuous one, $\alpha_2 = 1$. Moreover, for any reaction $y_r\to y_r'$ we have $\beta_r = 1$, and Assumptions 1 and 2 are fulfilled. Furthermore, the limiting rate functions defined in (9) are given by

$$\lambda_{A+B\to 2B}(x) := \lim_{N\to\infty} N^{-1}\lambda_{A+B\to 2B}^N([N^\alpha x]) = \lim_{N\to\infty} N^{-1}\kappa_1 x_1\lfloor Nx_2\rfloor = \kappa_1 x_1 x_2$$
$$\lambda_{B\to A}(x) := \lim_{N\to\infty} N^{-1}\lambda_{B\to A}^N([N^\alpha x]) = \lim_{N\to\infty} N^{-1}\kappa_2\lfloor Nx_2\rfloor = \kappa_2 x_2.$$

Turning to Example 1.2 we suppose the conserved quantity satisfies

$$X_1^N(0)+X_2^N(0)+X_3^N(0) = N,$$

where N is large, and that $X^N(0)$ is not far from the equilibrium q^N, which satisfies

$$q_1^N = \frac{\kappa_0\kappa_1\kappa_3}{\kappa_2+\kappa_3}\cdot\frac{q_2^N+1}{q_2^N-1},\quad q_3^N = \frac{\kappa_1}{\kappa_2+\kappa_3}q_2^N \quad\text{and}\quad q_1^N+q_2^N+q_3^N = N.$$

Therefore, for N large $X_1^N(0)$ will be near the value $q := \frac{\kappa_0\kappa_1\kappa_3}{\kappa_2+\kappa_3}$, while $X_2^N(0)$ and $X_3^N(0)$ go to infinity as $N\to\infty$. In this context, A is a discrete species, $\alpha_1 = 0$, and B and C are continuous species, $\alpha_2 = \alpha_3 = 1$.

Furthermore, $\beta_r = 1$ for each r, and Assumptions 1 and 2 hold. In this case, the limiting rate functions defined in (9) are given by

$$\lambda_{A+2B\to 3B}(x) := \lim_{N\to\infty} N^{-1}\kappa_0 \frac{x_1 N x_2 (N x_2 - 1)}{1 + N x_2} = \kappa_0 x_1 x_2,$$

and similarly,

$$\lambda_{B\to C}(x) := \kappa_1 x_2, \quad \lambda_{C\to B}(x) := \kappa_2 x_3, \quad \text{and} \quad \lambda_{C\to A}(x) = \kappa_3 x_3.$$

Returning to the general setting, let $\widetilde{\mathcal{R}}$ be the set of reactions whose source complex contain a continuous species, i.e.

$$\widetilde{\mathcal{R}} := \{y \to y' \in \mathcal{R} \,:\, \pi_c(y) \neq 0\}. \tag{12}$$

Due to (9), the reactions in $\widetilde{\mathcal{R}}$ have much higher rates than the other reactions, when N is large. Therefore, they give the major contribution to the dynamics of the stochastic system, and we focus on them.

We define two reduced systems, one being a projection onto the discrete species space and the other being the projection onto the continuous species space of the dynamics induced by the reactions in $\widetilde{\mathcal{R}}$. We begin by considering the projection onto the discrete species. Define

$$\begin{aligned}\pi_d(\mathcal{C}) &:= \{\pi_d(y) \,:\, y \in \mathcal{C}\},\\ \widetilde{\mathcal{R}}_d &:= \{\pi_d(y) \to \pi_d(y') \,:\, y \to y' \in \widetilde{\mathcal{R}} \text{ and } \pi_d(y) \neq \pi_d(y')\},\end{aligned}$$

and let $\mathcal{G}_d := (\mathcal{X}_d, \pi_d(\mathcal{C}), \widetilde{\mathcal{R}}_d)$ be the reaction network associated with the discrete species. For example, for both Example 1.1 and Example 1.2, the network associated with the discrete species is

$$A \rightleftharpoons 0.$$

Let $z_k \to z_k' \in \widetilde{\mathcal{R}}_d$. For any vector $w \in \mathbb{R}^{|\mathcal{X}_c|}_{\geq 0}$ we define the function $\lambda^w_{d,k} : \mathbb{R}^{|\mathcal{X}_d|} \to \mathbb{R}_{\geq 0}$ via

$$\lambda^w_{d,k}(v) := \sum_{y_r \to y_r' \in \widetilde{\mathcal{R}}^{d,k}} \lambda_r(v, w), \tag{13}$$

where $\widetilde{\mathcal{R}}^{d,k} := \{y_r \to y_r' \in \widetilde{\mathcal{R}} \,:\, \pi_d(y_r) = z_k \text{ and } \pi_d(y_r') = z_k'\}$. Let $\mathcal{K}^w_d$ be the kinetics defined by (13), and define $\mathscr{S}^w_d = (\mathcal{G}_d, \mathcal{K}^w_d)$. Note that the functions λ_r in (13) are the *limit* rate functions in (9). The sum in (13) is needed, as the cardinality of $\widetilde{\mathcal{R}}^{d,k}$ is not necessarily 1. Consider for example the following modification of Example 1.1:

$$\begin{gathered}A + B \xrightarrow{\kappa_1} 2B \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} A + 2B\\ B \xrightarrow{\kappa_2} A.\end{gathered}$$

The reactions $2B \to A + B$ and $B \to A$ collapse to the same reaction in $\widetilde{\mathcal{R}}_d$, and the same happens to $A + 2B \to 2B$ and $A + B \to 2B$. In this case, if w denotes the concentration of the species B, the system $\mathscr{S}^w_d$ is given by

$$A \underset{\kappa_2 w + \kappa_3 w^2}{\overset{\kappa_1 w + \kappa_4 w^2}{\rightleftharpoons}} 0.$$

We make another assumption on our models.

Assumption 3. The system $\mathscr{S}^w_d$ is endowed with stochastic mass action kinetics and it is complex balanced for any $w > 0$. We further require that $\mathcal{G}_d$ is irreducible.

Remark 3.2. If Assumption 2 is fulfilled, then each species of $\mathscr{S}^w_d$ appears with stoichiometric coefficient 1, which implies that for any choice of rate constants, the stochastic and the deterministic mass action kinetics coincide, due to their definition.

For the systems in Example 1.1 and in Example 1.2, $\mathscr{S}_d^w$ is given, respectively, by the stochastic mass action systems

$$A \underset{\kappa_2 w}{\overset{\kappa_1 w}{\rightleftharpoons}} 0 \quad \text{and} \quad A \underset{\kappa_3 w_2}{\overset{\kappa_0 w_1}{\rightleftharpoons}} 0, \tag{14}$$

where in the first system w represents the amount of species B and in the second system w_1 and w_2 represent the amounts of species B and C, respectively. In both cases, Assumption 3 holds, due to Theorem 2.2.

Thanks to Assumption 3 and Theorem 2.1, we know that for any positive w there is precisely one complex balanced equilibrium of the system $\mathscr{S}_d^w$, which we denote by q_d^w. For example, for the first system in (14) associated with Example 1.1 we have $q_d^w = \kappa_2/\kappa_1$, whereas for the second system in (14) associated with Example 1.2 we have $q_d^w = \kappa_3 w_2/\kappa_0 w_1$.

Assumption 3 is the last structural assumption we require for our main results. Now we impose some conditions ensuring that the systems are "well-behaved." Specifically, we want that the magnitude of the counts of each species is maintened constant in a compact interval of time. In particular, we want to rule out the possibility of a blow up or of a zeroing of the concentrations of the continuous species, and we want the counts of the discrete species to be bounded, in some sense.

We start by considering the projection onto the continuous species. Let

$$\begin{aligned}\pi_c(\mathcal{C}) &= \{\pi_c(y) \,:\, y \in \mathcal{C}\},\\ \widetilde{\mathcal{R}}_c &= \{\pi_c(y) \to \pi_c(y') \,:\, y \to y' \in \widetilde{\mathcal{R}} \text{ and } \pi_c(y) \neq \pi_c(y')\},\end{aligned}$$

and define $\mathcal{G}_c := (\mathcal{X}_c, \pi_c(\mathcal{C}), \widetilde{\mathcal{R}}_c)$. Consider the kth reaction $z_k \to z_k' \in \widetilde{\mathcal{R}}_c$, and define the function $\lambda_{c,k} : \mathbb{R}^{|\mathcal{X}_d|} \to \mathbb{R}_{\geq 0}$ via

$$\lambda_{c,k}(w) = \sum_{y_r \to y_r' \in \widetilde{\mathcal{R}}^{c,k}} \mathbb{1}_{\{w>0\}} \lambda_r(q_d^w, w), \tag{15}$$

where $\widetilde{\mathcal{R}}^{c,k} := \{y_r \to y_r' \in \widetilde{\mathcal{R}} \,:\, \pi_c(y_r) = z_k \text{ and } \pi_c(y_r') = z_k'\}$. Let $\mathcal{K}_c$ denote the kinetics defined by the above rate functions, and define $\mathscr{S}_c = (\mathcal{G}_c, \mathcal{K}_c)$. Finally, fix a finite time $T > 0$ and a point $X_0 \in \mathbb{R}_{>0}^{|\mathcal{X}|}$.

Assumption 4. Assume that the deterministic solution $z(t)$ of the system $\mathscr{S}_c$, with initial condition $\pi_c(X_0)$, exists for any $t \in [0, T]$. Moreover, assume that for all $t \in [0, T]$ we have $z(t) > 0$.

Consider Example 1.1. In this case, for any $w \in \mathbb{R}_{>0}$ we have $q_d^w = \kappa_2/\kappa_1$, and the system $\mathscr{S}_c$ is given by

$$0 \overset{\kappa_2}{\longleftarrow} B \overset{\kappa_2}{\longrightarrow} 2B. \tag{16}$$

Hence, in this case the deterministic solution $z(t)$ is constantly equal to $\pi_c(X_0)$.

Consider now Example 1.2. Here, for any $w \in \mathbb{R}_{>0}^2$, q_d^w is given by $\kappa_3 w_2/\kappa_0 w_1$. Therefore, the system $\mathscr{S}_c$ is

$$\begin{gathered}2B \longrightarrow 3B\\ B \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} C \overset{\kappa_3}{\longrightarrow} 0,\end{gathered}$$

with

$$\lambda_{2B \to 3B}(w) = \kappa_3 w_2.$$

When deterministically modeled, the dynamics of the system is equivalent to that of the deterministic mass action system

$$B \underset{\kappa_2+\kappa_3}{\overset{\kappa_1}{\rightleftharpoons}} C \tag{17}$$

and it can be easily shown that Assumption 4 holds, since $\pi_c(X_0) > 0$.

Assumption 5. There exists a locally bounded function $\psi\colon \mathbb{R}^{|\mathcal{X}_d|}_{\geq 0} \to \mathbb{R}_{\geq 1}$ satisfying

$$\text{(i)} \lim_{\|v\|\to\infty} \psi(v) = \infty \quad \text{and} \quad \text{(ii)} \sup_{N\in\mathbb{N}_{>0}} E\left[\int_0^T \psi(X^N_{\text{disc}}(s))ds\right] < \infty,$$

such that for any reaction $y_r \to y'_r \in \mathcal{R}$ and any compact set $\Gamma \subset \mathbb{R}^{|\mathcal{X}_c|}_{\geq 0}$

$$\text{(iii)} \sup_{w\in\Gamma} \limsup_{\|v\|\to\infty} \frac{\lambda_r(v,w)}{\psi(v)} = 0 \quad \text{and} \quad \text{(iv)} \sup_{\substack{w\in\Gamma \\ N\in\mathbb{N}_{>0}}} \limsup_{\|v\|\to\infty} \frac{\nu^N_r(v,w)}{\psi(v)} = 0,.$$

where ν^N_r is defined as in (10).

Remark 3.3. Assume that λ_r and ν^N_r have at most polynomial growth rate in v, for any reaction $y_r \to y'_r \in \mathcal{R}$ and any $N \in \mathbb{N}$. Let $p\colon \mathbb{R}^{|\mathcal{X}_d|} \to \mathbb{R}$ be a polynomial satisfying

$$\max_{y_r\to y'_r\in\mathcal{R}} \limsup_{\|v\|\to\infty} \frac{\nu^N_r(v,w)}{p(v)} = 0 \quad \text{and} \quad \max_{y_r\to y'_r\in\mathcal{R}} \limsup_{\|v\|\to\infty} \frac{\lambda_r(v,w)}{p(v)} = 0 \quad \text{for any } w \in \mathbb{R}^{|\mathcal{X}_c|}_{\geq 0},$$

and let d be the degree of p. Then, a candidate for ψ is

$$\psi(v) = 1 + \sum_{S_i\in\mathcal{X}_d} v_i^d.$$

With this choice, ψ automatically satisfies (i), (iii) and (iv).

We now state and prove our main result.

Theorem 3.1. *If Assumptions 1-5 hold, then*

$$N^{-1}X^N_{\text{cont}}(\cdot\wedge T) \underset{N\to\infty}{\Longrightarrow} z(\cdot\wedge T),$$

where z is defined as in Assumption 4 and $\Rightarrow$ denotes weak convergence in the Skorokhod topology, which in this case coincides with the uniform topology as the limit z is continuous. Moreover, if $\text{Pois}(q)$ denotes a product-form Poisson distribution with parameter q, then for any continuous function $g\colon \mathbb{R}^{|\mathcal{X}_d|}_{\geq 0} \to \mathbb{R}$ satisfying

$$\limsup_{\|v\|\to\infty} \frac{|g(v)|}{\psi(v)} = 0 \tag{18}$$

we have

$$\int_0^{\cdot\wedge T} \Big(g(X^N_{\text{disc}}(s)) - E[g(J^{z(s)})]\Big)ds \underset{N\to\infty}{\Longrightarrow} 0,$$

where $J^{z(s)} \sim \text{Pois}(q_d^{z(s)})$.

Proof. For the sake of simplicity, throughout the proof we will write t instead of $t\wedge T$, but it is always implicitly assumed that $t \in [0,T]$.

We follow the arguments of [12, 19], which rely on the techniques developed in [16]. We will first prove the theorem under the assumption that

$$\sup_N P\left(\sup_{t\in[0,T]} N^{-1}\|X^N_{\text{cont}}(t)\|_\infty > M\right) = 0, \tag{19}$$

for a certain constant M satisfying

$$\sup_{t\in[0,T]} \|z(t)\|_\infty < M.$$

This could happen if some conservation laws are present, or if we study the process up to the time when the concentration of a continuous species exceeds a given threshold. We will then drop the assumption (19).

Define the *occupation measures* Γ^N on $\mathbb{R}^{|\mathcal{X}_d|} \times [0,T]$ by

$$\Gamma^N(D \times [a,b]) = \int_a^b \mathbb{1}_D(X^N_{\text{disc}}(s))ds.$$

Note that

$$d\Gamma^N(v,s) = d\gamma^N_s(v)ds, \tag{20}$$

where $\gamma^N_s = \delta_{X^N_{\text{disc}}(s)}$, with δ_x denoting the usual Dirac measure on $\mathbb{R}^{|\mathcal{X}_d|}$. Part (ii) and (iv) in Assumption 5 together with Lemma 1.3 in [16] imply that the sequence of random measures Γ^N is tight. Let Γ be a weak limit point.

Consider the generator L^N for the process $N^{-\alpha}X^N(t)$, defined by

$$L^N f(x) = \sum_{y_r \to y'_r \in \mathcal{R}} \lambda^N_r(N^\alpha x)\Big(f\left(x + N^{-\alpha}(y'-y)\right) - f(x)\Big), \quad \text{for } N^\alpha x \in \mathbb{N}^{|\mathcal{X}|}.$$

From the generator L^N we can obtain two generators, one related to the limiting behaviour of the concentrations of the continuous species (whose changes take place at the time scale t) and the other one related to the discrete species (whose changes take place at the time scale $N^{-1}t$). For any function $h \in C^2_c(\mathbb{R}^{|\mathcal{X}_c|})$ and $x \in \mathbb{N}^{|\mathcal{X}_d|} \times \mathbb{R}^{|\mathcal{X}_c|}$, define

$$\begin{aligned} L_c h(x) &:= \lim_{N\to\infty} L^N(h\circ\pi_c)(N^{-\alpha}[N^\alpha x]) \\ &= \sum_{y_r \to y'_r \in \mathcal{R}} \lim_{N\to\infty} \lambda^N_r([N^\alpha x])\Big(h\Big(\pi_c(N^{-\alpha}[N^\alpha x]) + N^{-1}\pi_c(y'-y)\Big) - h\Big(\pi_c(N^{-\alpha}[N^\alpha x])\Big)\Big) \\ &= \sum_{y_r \to y'_r \in \widetilde{\mathcal{R}}} \lambda_r(x)\pi_c(y'-y)\cdot\nabla h(\pi_c(x)), \end{aligned}$$

where we made use of Assumption 1 to compute the limit, and $\cdot$ denotes the scalar product. Note that $L_c h \in C^1_c(\mathbb{R}^{|\mathcal{X}|})$ and that $L^N(h\circ\pi_c)(N^{-\alpha}[N^\alpha x])$ converges uniformly to $L_c h(x)$ on x. L_c can be interpreted as the generator of the limiting behaviour of the concentrations of the continuous species.

On the other hand, for any function $g \in C_c(\mathbb{N}^{|\mathcal{X}_d|})$ and $x \in \mathbb{N}^{|\mathcal{X}_d|} \times \mathbb{R}^{|\mathcal{X}_c|}$, define

$$\begin{aligned} L_d g(x) &:= \lim_{N\to\infty} N^{-1}L^N(g\circ\pi_d)(N^{-\alpha}[N^\alpha x]) \\ &= \sum_{y_r \to y'_r \in \mathcal{R}} \lim_{N\to\infty} N^{-1}\lambda^N_r([N^\alpha x])\Big(g\Big(\pi_d(x) + \pi_d(y'-y)\Big) - g\Big(\pi_d(x)\Big)\Big) \\ &= \sum_{y_r \to y'_r \in \widetilde{\mathcal{R}}} \lambda_r(x)\Big(g\Big(\pi_d(x) + \pi_d(y'-y)\Big) - g\Big(\pi_d(x)\Big)\Big). \end{aligned}$$

The convergence is uniform in x. L_d can be interpreted as the generator of the limiting behaviour of the discrete species. For any $w \in \mathbb{R}^{|\mathcal{X}_c|}_{\geq 0}$ we can define the operator L^w_d by

$$L^w_d g(v) := L_d g(v,w) \quad \forall v \in \mathbb{N}^{|\mathcal{X}_d|},$$

which corresponds to the generator of the system $\mathscr{S}^w_d$.

For any $h \in C^2_c(\mathbb{R}^{|\mathcal{X}_c|})$, the process

$$M^N_h(t) := h(N^{-1}X^N_{\text{cont}}(t)) - h(N^{-1}X^N_{\text{cont}}(0)) - \int_0^t L^N(h\circ\pi_c)(N^{-\alpha}X^N(s))ds \tag{21}$$

is a martingale. Let

$$\delta_h^N(t) := \int_0^t \left(L_c h - L^N(h \circ \pi_c)\right)(N^{-\alpha} X^N(s))ds.$$

By the uniform convergence of $L^N(h \circ \pi_c)(N^{-\alpha}[N^\alpha x])$ to $L_c h(x)$ with respect to x, we have that

$$\lim_{N\to\infty} E\left[\sup_{t\in[0,T]} |\delta_h^N(t)|\right] = 0. \tag{22}$$

Moreover, (ii) and (iii) in Assumption 5, together with (19), imply that

$$\sup_N E\left[\int_0^T |L_c h(N^{-\alpha} X^N(s))|ds\right] < \infty. \tag{23}$$

Since (21) is a martingale, tightness of the processes X^N_{cont} follows from (19), (22) and (23) by Theorems 3.9.1 and 3.9.4 in [8]. Let (W, Γ) be a weak limit of $(X^N_{\text{cont}}, \Gamma^N)$. By the same arguments as in the proof of Theorem 2.1 in [16], we have that

$$M_h(t) := h(W(t)) - h(W(0)) - \int_{\mathbb{R}^{|\mathcal{X}_d|}\times[0,t]} L_c h(v, W(s))d\Gamma(v, s)$$

is a martingale. On the other hand, for any $g \in C_c^2(\mathbb{R}^{|\mathcal{X}_d|})$, the process

$$\widehat{M}_g^N(t) := N^{-1}\left[g(X^N_{\text{disc}}(t)) - g(X^N_{\text{disc}}(0)) - \int_0^t L^N(g \circ \pi_d)(N^{-\alpha} X^N(s))ds\right]$$

is also a martingale. Since the function g is bounded and by using uniform convergence of $N^{-1}L^N(g \circ \pi_d)$ to $L_d g$, we have that a weak limit point for $\widehat{M}_g^N$ is given by

$$\widehat{M}_g(t) := -\int_{\mathbb{R}^{|\mathcal{X}_d|}\times[0,t]} L_d g(v, W(s))d\Gamma(v, s),$$

which is therefore a martingale. By (20) we have

$$d\Gamma(v, s) = d\gamma_s(v)ds, \tag{24}$$

for a family of measures γ_s. Therefore, $\widehat{M}_g(t)$ is continuous and for any $t_1 < t_2$

$$\int_{\mathbb{R}^{|\mathcal{X}_d|}\times[t_1,t_2]} |L_d g(v, W(s))|\, d\Gamma(v, s) \le (t_2 - t_1) \sup_{x\in\mathbb{R}^{\mathcal{X}}} |L_d g(x)|,$$

which implies that $\widehat{M}_g(t)$ has finite variation paths. This in turn implies that $\widehat{M}_g(t)$ is constantly equal to $\widehat{M}_g(0) = 0$ for any $t \in [0, T]$ with probability one. Therefore, almost surely, for almost every $s \in [0, T]$

$$\int_{\mathbb{R}^{|\mathcal{X}_d|}} L_d^{W(s)} g(v)d\gamma_s(v) = 0,$$

where γ_s is as in (24). Since $C_c^2(\mathbb{R}^{|\mathcal{X}|})$ is separable, we have that, for almost every $s \in [0, T]$,

$$\int_{\mathbb{R}^{|\mathcal{X}_d|}} L_d^{W(s)} g(v)d\gamma_s(v) = 0, \quad \forall g \in C_c^2(\mathbb{R}^{|\mathcal{X}_d|}).$$

Thus, for almost every s such that $W(s) > 0$ with probability one, the measure γ_s is equal to the *unique* stationary distribution of the system $\mathscr{S}_d^{W(s)}$. Due to Assumption 3 and by [2], it corresponds to the product form Poisson distribution $\text{Pois}(q_d^{W(s)})$. Therefore we can write

$$\gamma_s = \mathbb{1}_{\{W(s)>0\}}\text{Pois}(q_d^{W(s)}) + (1 - \mathbb{1}_{\{W(s)>0\}})\gamma_s.$$

The weak limit process $W(t)$ is a solution to the *martingale problem*

$$\begin{aligned}
M_h(t) =& h(W(t)) - h(W(0)) - \int_0^t \int_{\mathbb{R}^{|\mathcal{X}_d|}} L_c h(v, W(s)) d\gamma_s(v) ds \\
=& h(W(t)) - h(W(0)) - \sum_{y_r \to y_r' \in \widetilde{\mathcal{R}}} \int_0^t \mathbb{1}_{\{W(s)>0\}} \lambda_r(q_d^{W(s)}, W(s)) \pi_c(y' - y) \cdot \nabla h(W(s)) ds \\
& - \int_0^t (1 - \mathbb{1}_{\{W(s)>0\}}) \int_{\mathbb{R}^{|\mathcal{X}_d|}} L_c h(v, W(s)) d\gamma_s(v) ds,
\end{aligned}$$

where we used Assumption 2 in the last equality. By Assumption 4, $W(t)$ is uniquely determined by the *unique* solution to

$$\begin{aligned}
W(0) =& \pi_c(X_0) \\
W(t) =& W(0) + \sum_{y_r \to y_r' \in \widetilde{\mathcal{R}}} \int_0^t \mathbb{1}_{\{W(s)>0\}} \lambda_r(q_d^{W(s)}, W(s)) \pi_c(y' - y) ds + \\
& + \int_0^t (1 - \mathbb{1}_{\{W(s)>0\}}) \int_{\mathbb{R}^{|\mathcal{X}_d|}} L_c id(v, W(s)) d\gamma_s(v) ds \\
=& W(0) + \sum_{y_r \to y_r' \in \widetilde{\mathcal{R}}} \int_0^t \lambda_r(q_d^{W(s)}, W(s)) \pi_c(y' - y) ds,
\end{aligned}$$

which is given by $z(t)$ (see [23, 17]). The first part of the Proposition is therefore proved. The second part follows from Lemma 2.9 in [13].

To prove the Theorem holds without assuming (19), fix a constant M with the property

$$M - \sup_{t \in [0,T]} \|z(t)\|_\infty > \delta,$$

for some $\delta > 0$, and consider the stopping time

$$\tau^N := \inf\{t \in [0,T] \,:\, N^{-1} \|X_{\text{cont}}^N\|_\infty > M\}.$$

Note that $N^{-1} X_{\text{cont}}(\tau^N)$ is also uniformly bounded in N, as

$$N^{-1} \|X_{\text{cont}}^N(\tau^N)\|_\infty \leq M + \max_r \{\|\xi_r\|_\infty\} := M'.$$

Therefore, (19) holds up to time τ^N with the constant M', which means that for any ε there exists N_ε such that for any $N > N_\varepsilon$

$$P\left(\sup_{t \in [0, \tau^N]} \|X_{\text{cont}}^N(t) - z(t)\| > \delta \right) < \varepsilon.$$

It follows that for N large enough $P(\tau^N < T) < \varepsilon$. The proof is concluded by the arbitrariness of ε, since for N large enough

$$P\left(\sup_{t \in [0,T]} \|X_{\text{cont}}^N(t) - z(t)\| > \eta \right) \leq P\left(\sup_{t \in [0, \tau^N]} \|X_{\text{cont}}^N(t) - z(t)\| > \eta \right) + \varepsilon,$$

$$P\left(\sup_{t \in [0,T]} \int_0^t \Big(g(X_{\text{disc}}^N(s)) - E[g(J^{z(s)})] \Big) ds > \eta \right) \leq P\left(\sup_{t \in [0, \tau^N]} \int_0^t \Big(g(X_{\text{disc}}^N(s)) - E[g(J^{z(s)})] \Big) ds > \eta \right) + \varepsilon.$$

□

Remark 3.4. If Assumption 2 is not fulfilled, Theorem 3.1 still holds if (15) is substituted with

$$\lambda_{c,r}(w) = \sum_{\substack{y_r \to y_r' \in \mathcal{R}^{c,k} \\ \pi_c(y) \neq 0}} \mathbb{1}_{\{w>0\}} E[\lambda_r(J^w, w)],$$

where $J^w \sim \text{Pois}(q_d^w)$. The proof of Theorem 3.1 also covers this more general case, provided that every occurence of $\lambda_r(q_d^w, w)$ is substituted with $E[\lambda_r(J^w, w)]$.

Remark 3.5. If Assumptions 2 and 3 are not satisfied, but we know that for any $w \in \mathbb{R}_{>0}^{|\mathcal{X}_c|}$ the stochastic system $\mathscr{S}_d^w$ possesses a unique stationary distribution μ^w with

$$\overline{\lambda}_r(w) := \sum_{v \in \mathbb{N}^{\mathcal{X}_d}} \lambda_r(v, w)\mu^w(v) < \infty \quad \text{for every } y_r \to y_r' \in \mathcal{R},$$

then Theorem 3.1 still holds, provided that every occurrence of $\lambda_r(q_d^w, w)$ is substituted with $\overline{\lambda}_r(w)$. The proof of Theorem 3.1, with small changes, also covers this generalization. In regard to this broader setting, see also the results in [12, 19].

In some cases Assumption 5 can be difficult to check, even if it seems natural for the analysed system. For this reason, we state here a corollary of Theorem 3.1 concerning a particular case for which Assumption 5 is automatically satisfied.

Corollary 3.2. *Assume Assumptions 1-4 hold. Assume also that*

$$N^{-\beta_r}\lambda_r^N(v, [Nw]) \leq \widehat{h}_r(w)v^{\pi_d(y_r)} \quad \textit{for any } y_r \to y_r' \in \mathcal{R}, \tag{25}$$

for some continuous positive functions $\widehat{h}_r \colon \mathbb{R}^{|\mathcal{X}_c|} \to \mathbb{R}$*, and that*

$$N^{-\beta_r}\lambda_r^N(v, [Nw]) \geq h_r(w)w^{\pi_c(y_r)}v^{\pi_d(y_r)} \quad \textit{for any } y_r \to y_r' \in \widetilde{\mathcal{R}}, \tag{26}$$

for some continuous positive functions $h_r \colon \mathbb{R}^{|\mathcal{X}_c|} \to \mathbb{R}$*. Furthermore, assume that in the support of any complex* $y \in \mathcal{C}$ *at most one discrete species appears. Then, for any continuous function* $g \colon \mathbb{R}_{\geq 0}^{|\mathcal{X}_d|} \to \mathbb{R}$ *with at most polynomial growth rate we have*

$$\int_0^{\cdot \wedge T} \left(g(X_{\text{disc}}^N(s)) - E[g(J^{z(s)})] \right) ds \xRightarrow[N\to\infty]{} 0,$$

where $J^w \sim \text{Pois}(q_d^w)$ *and* $z(t)$ *is as in Assumption 4.*

Proof. By Assumption 4, we can choose two positive constants $m < M$ such that

$$\inf_{S_i \in \mathcal{X}_c, t \in [0,T]} Z_i(t) > m \quad \text{and} \quad \sup_{S_i \in \mathcal{X}_c, t \in [0,T]} Z_i(t) < M. \tag{27}$$

For any $N \in \mathbb{N}_{>0}$, consider the function $\pi^N : \mathbb{R}^{|\mathcal{X}|} \to \mathbb{R}^{|\mathcal{X}|}$ defined by

$$(\pi^N(x))_i := \begin{cases} (Nm \vee x_i) \wedge NM & \text{if } S_i \in \mathcal{X}_c \\ x_i & \text{otherwise.} \end{cases}$$

Consider the modified family of kinetics $\overline{\mathcal{K}}^N$ defined by

$$\overline{\lambda}_r^N(x) := \lambda_r^N\left(\pi^N(x)\right) \quad \text{for } x \in \mathbb{N}^{|\mathcal{X}|}.$$

We have

$$\lim_{N\to\infty} N^{-\beta_r}\overline{\lambda}_r^N([N^\alpha x]) = \overline{\lambda}_r(x) := \lambda_r\left(\pi^1(x)\right),$$

where the limit is uniform on compact sets. Furthermore, $\overline{\lambda}_r$ is locally Lipschitz. Let $\overline{\mathcal{K}}$ be the kinetics defined by the functions $\overline{\lambda}_r$. Our first aim is to prove that for a suitable choice of ψ, Assumption 5 holds for the modified kinetics. We then apply Theorem 3.1 to the reaction systems $(\mathcal{G}, \overline{\mathcal{K}}^N)$. Let $\overline{X}^N(t)$ denote the stochastic process associated with $(\mathcal{G}, \overline{\mathcal{K}}^N)$. Define

$$\sigma^N(t) := \sum_{S_i \in \mathcal{X}_d} \overline{X}_i^N(t) = \sum_{S_i \in \mathcal{X}_d} X_i^N(t) \quad \text{and} \quad \Delta_r := \sum_{S_i \in \mathcal{X}_d} \xi_{ri}.$$

Since the complexes y are non-negative vectors, we have

$$\Delta_r = \sum_{S_i \in \mathcal{X}_d} (y'_{ri} - y_{ri}) = \|\pi_d(y'_r)\|_1 - \|\pi_d(y_r)\|_1.$$

By hypothesis and by Assumption 2, for any complex $y \in \mathcal{C}$ we have $\|\pi_d(y)\|_1 \leq 1$, which implies that $-1 \leq \Delta_r \leq 1$ for any $y_r \to y'_r \in \mathcal{R}$. Moreover, we have

$$\begin{array}{rcl} \Delta_r = 1 & \implies & \|\pi_d(y_r)\|_1 = 0 \\ \Delta_r = -1 & \implies & \|\pi_d(y_r)\|_1 = 1. \end{array} \tag{28}$$

Furthermore,

$$\begin{aligned} \sigma^N(t) &= \sigma^N(0) + \sum_{y_r \to y'_r \in \mathcal{R}} \Delta_r Y_r \left(\int_0^t \overline{\lambda}_r^N \left(\overline{X}^N(s) \right) ds \right) \\ &= \sigma^N(0) + \sum_{\substack{y_r \to y'_r \in \mathcal{R} \\ \Delta_r = 1}} Y_r \left(\int_0^t \overline{\lambda}_r^N \left(\overline{X}^N(s) \right) ds \right) - \sum_{\substack{y_r \to y'_r \in \mathcal{R} \\ \Delta_r = -1}} Y_r \left(\int_0^t \overline{\lambda}_r^N \left(\overline{X}^N(s) \right) ds \right) \end{aligned} \tag{29}$$

Define

$$M^* := \max_{y_r \to y'_r \in \mathcal{R}} \max_{m \leq w \leq M} \widehat{h}_r(w) \qquad \text{and} \qquad m^* := \min_{y_r \to y'_r \in \mathcal{R}} \min_{m \leq w \leq M} h_r(w) w^{\pi_c(y_r)},$$

which are both positive constants. By (28) and (25), whenever $\Delta_r = 1$ we have

$$\overline{\lambda}_r^N(v, Nw) \leq NM^*$$

On the other hand, by (26), if $y_r \to y'_r \in \widetilde{\mathcal{R}}$ and $\Delta_r = -1$ then

$$\overline{\lambda}_r^N(v, Nw) \geq N^{\beta_r} m^* v_i \quad \text{for some } S_i \in \mathcal{X}_d.$$

By Assumption 3 the system $\mathscr{S}_d^w$ is complex balanced, which implies that it is weakly reversible by Theorem 2.1. In particular, for any $S_i \in \mathcal{X}_c$ we can choose a reaction $y_{r(i)} \to y'_{r(i)} \in \widetilde{\mathcal{R}}$ such that $\xi_{r(i)i} = -1$. We have that

$$\sum_{\substack{y_r \to y'_r \in \mathcal{R} \\ \Delta_r = -1}} Y_r \left(\int_0^t \overline{\lambda}_r^N \left(\overline{X}^N(s) \right) ds \right) \geq \sum_{S_i \in \mathcal{X}_d} Y_{r(i)} \left(\int_0^t N m^* \overline{X}_i^N(s) ds \right).$$

Then, from (29) it follows that

$$\begin{aligned} \sigma^N(t) &\leq \sigma^N(0) + \sum_{y_r \to y'_r \in \mathcal{R}} Y_r \left(\int_0^t NM^* ds \right) - \sum_{S_i \in \mathcal{X}_d} Y_{r(i)} \left(\int_0^t N m^* \overline{X}_i^N(s) ds \right) \\ &\sim \sigma^N(0) + Y' \Big(|\mathcal{R}| \cdot NM^* t \Big) - Y'' \left(\int_0^t N m^* \sigma^N(s) ds \right), \end{aligned}$$

for two i.i.d. unit-rate Poisson processes Y' and Y''. This means that if we define the birth-death process

$$B(t) := \sigma^N(0) + Y'\Big(|\mathcal{R}| \cdot M^* t\Big) - Y''\left(\int_0^t m^* B(s) ds\right),$$

we have that

$$E[\sigma^N(t)^n] \leq E[B(Nt)^n].$$

Moreover, for any $n \in \mathbb{N}$

$$\sup_{\substack{t\in[0,T]\\ N\in\mathbb{N}_{>0}}} E[\sigma^N(t)^n] \leq \sup_{t\in[0,\infty)} E[B(t)^n] < \infty. \tag{30}$$

For any $n > 1$, we define $\psi_n \colon \mathbb{R}_{\geq 0}^{|\mathcal{X}_d|} \to \mathbb{R}_{\geq 1}$ via

$$\psi_n(v) = 1 + \sum_{S_i \in \mathcal{X}_d} v_i^n.$$

Due to 25, λ_r^N has at most polynomial growth rate in v, for any reaction $y_r \to y_r' \in \mathcal{R}$ and any $N \in \mathbb{N}$. By (9), 25 also implies that the rate functions λ_r have at most polynomial growth rate. This, due to (10), in turn implies that ν_r^N has at most polynomial growth rate in v, for any reaction $y_r \to y_r' \in \mathcal{R}$ and any $N \in \mathbb{N}$. By Remark 3.3, if n is large enough, ψ_n satisfies (i), (iii) and (iv) in Assumption 5. Moreover,

$$E\left[\int_0^T \psi_n(\overline{X}^N_{\text{disc}}(s)) ds\right] \leq E\left[\int_0^T \sigma^N(s)^n ds\right],$$

hence, due to Fubini's Theorem and (30), part (ii) in Assumption 5 is verified, too. Assumptions 1-3 also hold for the systems with modified rates. Moreover, due to (27), the solution of the deterministic system $(\mathcal{G}_c, \overline{\mathcal{K}}_c)$ coincide with Z, the solution of the deterministic system $\mathscr{S}_c$. Therefore, Assumption 4 is satisfied as well and we can apply Theorem 3.1 to the modified reaction systems $(\mathcal{G}_c, \overline{\mathcal{K}}^N)$. We have

$$N^{-1}\overline{X}^N_{\text{cont}}(\cdot \wedge T) \xRightarrow[N\to\infty]{} z(\cdot \wedge T), \tag{31}$$

and, since by definition any function $g \colon \mathbb{R}_{\geq 0}^{|\mathcal{X}_d|} \to \mathbb{R}$ with at most polynomial growth rate satisfies

$$\limsup_{\|v\|\to\infty} \frac{|g(v)|}{\psi_n(v)} = 0$$

for n large enough, we have

$$\int_0^{\cdot\wedge T} \Big(g(\overline{X}^N_{\text{disc}}(s)) - E[g(J^{z(s)})]\Big) ds \xRightarrow[N\to\infty]{} 0.$$

The proof is completed by noting that if the path of $\overline{X}^N_{\text{disc}}$ is different from the path of X^N_{disc}, then we have

$$\inf_{S_i\in\mathcal{X}_c, t\in[0,T]} \overline{X}_i^N(t) \leq m \quad \text{or} \quad \sup_{S_i\in\mathcal{X}_c, t\in[0,T]} \overline{X}_i(t) \geq M.$$

However, by (27) and (31), we have

$$P\left(\inf_{S_i\in\mathcal{X}_c, t\in[0,T]} \overline{X}_i^N(t) \leq m \quad \text{or} \quad \sup_{S_i\in\mathcal{X}_c, t\in[0,T]} \overline{X}_i(t) \geq M\right) \xrightarrow[N\to\infty]{} 0.$$

□

Remark 3.6. If $\mathcal{K}^N$ is mass action and the rate constants are rescaled according to (11), then (25) and (26) are automatically satisfied. Indeed, if Assumption 2 holds, then

$$N^{-\beta_r}\lambda_r^N(v,[Nw]) = (\kappa_r + \varepsilon_N^1)(w^{\pi_c(y_r)} + \varepsilon_N^2(w))v^{\pi_d(y_r)},$$

for some sequence ε_N^1 converging to zero, and some function $\varepsilon_N^2(w)$ converging to zero uniformly on w. In this regard, see Remark 3.1.

4 ACR setting

We turn to the ACR setting and start with the formal definition of *absolute concentration robustness* (ACR).

Definition 4.1. Let $\mathscr{S} = (\mathcal{G}, \mathcal{K})$ be a reaction system. We say that a species S_i possesses *absolute concentration robustness* (ACR) in $\mathscr{S}$ if for any two positive equilibria q, q' of the system $\mathscr{S}$, we have $q_i = q_i'$. In this case, the species S_i is called an *ACR species* and, if a positive equilibrium q exists, q_i is called an *ACR value*. If a system $\mathscr{S}$ possess a non empty set of ACR species, we call it an ACR system.

Consider a reaction system $\mathscr{S}$ that has no equilibria or a unique equilibrium. According to Definition 4.1, all the species of $\mathscr{S}$ are ACR species, however in these cases the ACR property is not particularly meaningful.

Definition 4.2. We say that a system $\mathscr{S}$ is a *non-degenerate* ACR system if it is an ACR system and possesses at least two positive equilibria. If an ACR system exhibits less than two positive equilibria, we call it a *degenerate* ACR system.

We will focus on non-degenerate ACR systems. Note that in such systems not all species can be ACR species.

In non-degenerate ACR systems, the ACR species maintain their equilibrium concentration regardless the total amount of molecules present in the system. Our goal is to study the behaviour of the system when the abundance of species that do not exhibit ACR tends to infinity. It seems therefore natural to use the setting developed in the Section 3 and let the ACR species coincide with the discrete species. We further assume that the rate functions are rescaled consistently with the hypotheses of Section 3, such that (9) holds uniformly on compact sets.

In order to study the limit behaviour of ACR systems, we state some corollaries of Theorem 3.1, assuming the next assumption is satisfied.

Assumption 6. Consider the limit rate functions λ_r fulfilling (9), and the subset of reactions $\widetilde{\mathcal{R}}$ defined in (12). Let $\widetilde{\mathscr{S}} := (\mathcal{X}, \mathcal{C}, \widetilde{\mathcal{R}}, \widetilde{\mathcal{K}})$, where the kinetics $\widetilde{\mathcal{K}} \subseteq \mathcal{K}$ is the set of functions λ_r with $y_r \to y_r' \in \widetilde{\mathcal{R}}$. We assume that $\mathscr{S}$ is a non-degenerate ACR system, and that at least one of the ACR species is a discrete species.

Consider Example 1.1. The system $\widetilde{\mathscr{S}}$ coincides with the system introduced in the example itself, namely (1), and it is a non-degenerate ACR system. On the other hand, in Example 1.2 the system $\widetilde{\mathscr{S}}$ is given by

$$\begin{aligned} A + 2B &\longrightarrow 3B \\ B &\underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} C \xrightarrow{\kappa_3} A, \end{aligned}$$

with

$$\lambda_{A+2B\to 3B}(x) = \kappa_0 x_1 x_2.$$

Therefore, the dynamics of $\widetilde{\mathscr{S}}$, if deterministically modeled, coincide with that of the mass action system

$$\begin{aligned} A + B &\xrightarrow{\kappa_0} 2B \\ B &\underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} C \xrightarrow{\kappa_3} A, \end{aligned} \tag{32}$$

which is a non-degenerate ACR system with equilibria determined by the equations

$$x_1 = \frac{\kappa_1\kappa_3}{\kappa_0(\kappa_2+\kappa_3)} \quad \text{and} \quad x_2 = \frac{\kappa_2+\kappa_3}{\kappa_1}x_3.$$

When Assumption 6 is fulfilled, denote by $\mathcal{X}_{\text{ACR}}$ the set of discrete ACR species. Let $\pi_{\text{ACR}}\colon \mathbb{R}^{|\mathcal{X}_d|} \to \mathbb{R}^{|\mathcal{X}_{\text{ACR}}|}$ be the projection onto the species of $\mathcal{X}_{\text{ACR}}$, and let $X_{\text{ACR}}(t) = \pi_{\text{ACR}}(X_d(t))$. Finally, let q be the vector of the ACR values for the species in $\mathcal{X}_{\text{ACR}}$, and let

$$J \sim \text{Pois}(q). \tag{33}$$

Corollary 4.1. *Suppose that Assumptions 1-4 and 6 hold. Moreover, assume that (25) and (26) hold, and that in the support of any complex $y \in \mathcal{C}$ at most one discrete species appears. Let $\pi_c(X_0)$ be a positive equilibrium point for $\mathscr{S}_c$, and let J be as in (33). Then, for any continuous function $\widehat{g}\colon \mathbb{R}^{|\mathcal{X}_{\text{ACR}}|}_{\geq 0} \to \mathbb{R}$ with at most polynomial growth rate we have*

$$\int_0^{\cdot\wedge T} \Big(\widehat{g}(X^N_{\text{ACR}}(s)) - E[\widehat{g}(J)]\Big)ds \Longrightarrow_{N\to\infty} 0. \tag{34}$$

In particular,

$$\int_0^{\cdot\wedge T} \Big(X^N_{\text{ACR}}(s) - q\Big)ds \Longrightarrow_{N\to\infty} 0. \tag{35}$$

Proof. Since $\pi_c(X_0)$ is an equilibrium point for $\mathscr{S}_c$, we have $z(t) = \pi_c(X_0)$ for any $t \in [0,T]$. Moreover, by definition q_d^w is the complex balancing equilibrium point of the system $\mathscr{S}_d^w$. By Assumption 2 and Remark 3.2, $(q_d^{\pi_c(X_0)}, \pi_c(X_0))$ is a positive equilibrium point for $\widetilde{\mathscr{S}}$. Hence, $\pi_{\text{ACR}}\left(q_d^{\pi_c(X_0)}\right) = q$ and (34) follows from Corollary 3.2, applied to the function $g = \widehat{g} \circ \pi_{\text{ACR}}$. Therefore, for any $S_i \in \mathcal{X}_{\text{ACR}}$ we can apply the result to the function $x_i - q_i$ to obtain

$$\int_0^{\cdot\wedge T} \Big(X^N_i(s) - q_i\Big)ds \Longrightarrow_{N\to\infty} 0,$$

which concludes the proof. □

Example 4.1. Consider the reaction network in Example 1. $\mathscr{S}_c$ is given by (16), for which any non-negative real point is an equilibrium point. We choose the sequence of starting points $X^N(0) = (a_N, N(b+\varepsilon_N))$, where $(a_N)_{N\in\mathbb{N}}$ is a bounded sequence of natural numbers, b is a positive real number and $(\varepsilon_N)_{N\in\mathbb{N}}$ is a sequence of real numbers tending to zero. Therefore, the hypotheses of Corollary 4.1 are fulfilled for any positive T. In this case there is only one ACR species, namely A, and it is the only discrete species. Hence, for any continuous function $\widehat{g}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ with at most polynomial growth rate we have

$$\int_0^{\cdot\wedge T} \Big(\widehat{g}(X^N_1(s)) - E[\widehat{g}(J)]\Big)ds \Longrightarrow_{N\to\infty} 0,$$

where $X^N_1(s)$ is the counts for the species A at time s, for the starting point $X^N(0)$. This means that on average, the counts of the species A tend to follow a Poisson distribution with mean given by the ACR value $q = \kappa_2/\kappa_1$. □

Corollary 4.2. *Suppose that Assumptions 1-4 hold for any $T > 0$, and that Assumption 6 is verified. Moreover, assume that (25) and (26) hold, and that in the support of any complex $y \in \mathcal{C}$ at most one discrete species appears. Assume that $\pi_c(X_0)$ is in the basin of attraction of an equilibrium point of $\mathscr{S}_c$, and let J be as in (33). Then, for any continuous function $\widehat{g}\colon \mathbb{R}^{|\mathcal{X}_{\text{ACR}}|}_{\geq 0} \to \mathbb{R}$ with at most polynomial growth rate and for any $\varepsilon > 0$, there exists $H_\varepsilon > 0$ such that, for any $t_1, t_2 \in \mathbb{R}$ with $t_2 > t_1 > H_\varepsilon$,*

$$\lim_{N\to\infty} P\left(\sup_{t\in[t_1,t_2]}\left|\int_{t_1}^{t}\Big(\widehat{g}(X^N_{\text{ACR}}(s)) - E[\widehat{g}(J)]\Big)ds\right| > \varepsilon(1+t_2-t_1)\right) = 0. \tag{36}$$

In particular,

$$\lim_{N\to\infty} P\left(\sup_{t\in[t_1,t_2]}\left|\int_{t_1}^{t}\Big(X^N_{\mathrm{ACR}}(s)-q\Big)ds\right|>\varepsilon(1+t_2-t_1)\right)=0 \tag{37}$$

Proof. For the sake of simplicity, throughout this proof w will denote a vector varying in $\mathbb{R}^{|\mathcal{X}_c|}_{>0}$, even if not explicitly stated.

First, recall that, regardless the value $w>0$, the complex balanced equilibrium q_d^w is the unique solution of a system of multivariate polynomials, and as such, it is continuous in the coefficients of the polynomials [22, Chapter 8]. In particular, it is a continuous function of $w>0$.

Consider a sequence of vectors $(w_n)_{n\in\mathbb{N}}\subset\mathbb{R}^{|\mathcal{X}_c|}_{>0}$ converging to $w^*>0$. Therefore, the sequence $(q_d^{w_n})_{n\in\mathbb{N}}$ converges to $q_d^{w^*}$ and by Lebesgue's Dominated Convergence Theorem we have that

$$E[\widehat{g}(\pi_{\mathrm{ACR}}(J^{w_n}))]\xrightarrow[n\to\infty]{}E[\widehat{g}(\pi_{\mathrm{ACR}}(J^{w^*}))],$$

where $J^w\sim\mathrm{Pois}(q_d^w)$. This implies that $E[\widehat{g}(\pi_{\mathrm{ACR}}(J^w))]$ is a continuous function of w.

Let w^* be the equilibrium point of $\mathscr{S}_c$ whose basin of attraction $\pi_c(X_0)$ belongs. Since $(q_d^{w^*},w^*)$ is an equilibrium point of the system $\widetilde{\mathscr{S}}$ considered in Assumption 6, we have that $\pi_{\mathrm{ACR}}(q_d^{w^*})=q$. In particular, $\pi_{\mathrm{ACR}}(J^{w^*})\sim J$. Let δ_ε be a positive real number such that

$$\left|E[\widehat{g}(\pi_{\mathrm{ACR}}(J^w))]-E[\widehat{g}(J)]\right|<\varepsilon\quad\text{for any }\|w-w^*\|<\delta_\varepsilon.$$

Finally, let $H_\varepsilon>0$ be such that

$$\|z(t)-w^*\|<\delta_\varepsilon\quad\text{for any } t>H_\varepsilon.$$

Fix $t_1,t_2\in\mathbb{R}$ with $t_2>t_1>H_\varepsilon$. Since by hypothesis Assumptions 1-4 hold for any fixed $T>0$, due to Corollary 3.2 and by choosing $g=\widehat{g}\circ\pi_{\mathrm{ACR}}$ we have that

$$\lim_{N\to\infty} P\left(\sup_{t\in[t_1,t_2]}\left|\int_{t_1}^{t}\Big(\widehat{g}(X^N_{\mathrm{ACR}}(s))-E[\widehat{g}(\pi_{\mathrm{ACR}}(J^{z(s)}))]\Big)ds\right|>\varepsilon\right)=0.$$

Since $t_2>t_1>H_\varepsilon$, we have

$$\begin{aligned}
&\left|\int_{t_1}^{t}\Big(\widehat{g}(X^N_{\mathrm{ACR}}(s))-E[\widehat{g}(\pi_{\mathrm{ACR}}(J^{z(s)}))]\Big)ds\right|\\
&\qquad=\left|\int_{t_1}^{t}\Big(\widehat{g}(X^N_{\mathrm{ACR}}(s))-E[\widehat{g}(J)]\Big)ds+\int_{t_1}^{t}\Big(E[\widehat{g}(J)]-E[\widehat{g}(\pi_{\mathrm{ACR}}(J^{z(s)}))]\Big)ds\right|\\
&\qquad\geq\left|\int_{t_1}^{t}\Big(\widehat{g}(X^N_{\mathrm{ACR}}(s))-E[\widehat{g}(J)]\Big)ds\right|-\left|\int_{t_1}^{t}\Big(E[\widehat{g}(J)]-E[\widehat{g}(\pi_{\mathrm{ACR}}(J^{z(s)}))]\Big)ds\right|\\
&\qquad\geq\left|\int_{t_1}^{t}\Big(\widehat{g}(X^N_{\mathrm{ACR}}(s))-E[\widehat{g}(J)]\Big)ds\right|-\varepsilon(t_2-t_1),
\end{aligned}$$

and (36) follows. By applying the result to the functions x_i-q_i for any $S_i\in\mathcal{X}_d$, we also obtain (37). □

Example 4.2. Consider Example 1.2. Let $X^N(0)=(a_N,N(b+\varepsilon_N^1),N(c+\varepsilon_N^2))$, where $(a_N)_{N\in\mathbb{N}}$ is a bounded sequence of natural numbers, b,c are positive real numbers and $(\varepsilon_N^1)_{N\in\mathbb{N}}$ and $(\varepsilon_N^2)_{N\in\mathbb{N}}$ are two sequences of real numbers tending to zero. The continuous system $\mathscr{S}_c$, when deterministically modeled, is equivalent to (17). Therefore the hypotheses of Corollary 4.2 are fulfilled, since any $(b,c)\in\mathbb{N}^2_{>0}$ is in the basin of attraction of

$$w^*=\left(\frac{(\kappa_2+\kappa_3)(b+c)}{\kappa_1+\kappa_2+\kappa_3},\frac{\kappa_1(b+c)}{\kappa_1+\kappa_2+\kappa_3}\right).$$

Let q be the ACR value for A in the system (32). Therefore, after some time the counts of A are approximately distributed, on average, as a Poisson random variable with mean q, in the sense of (36). □

We conclude this section by an example that does not fit in our theory, but it is still tractable.

Example 4.3. Consider the stochastic mass action system

$$2A + B \xrightarrow{\kappa_1} 3B$$
$$B \xrightarrow{\kappa_2} A$$

Due to [20], we know that A possesses ACR and the system is a non-degenerate ACR system. The discrete system $\mathscr{S}_d^w$ is given by

$$2A \xrightarrow{\kappa_1 w} 0 \xrightarrow{\kappa_2 w} A.$$

Therefore, Assumption 3 does not hold, and by [paper coming soon!] we know that the $\mathscr{S}_d^w$, stochastically modeled, cannot exhibit a Poisson stationary distribution. However, it can be proven that a stationary distribution μ^w exists (for example by using Tweedie Recurrence Criterion [citation]). By following the proofs of the results in this paper, we can still argue that, if $X^N(0) = (a, bN)$ and $J^w \sim \mu^w$, for any continuous function $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ with at most polynomial growth rate we have

$$\int_0^{\cdot \wedge T} \Big(g(X_1^N(s)) - E[g(J^{z(s)})] \Big) ds \Longrightarrow_{N\to\infty} 0. \tag{38}$$

In this regard, see Remark 3.5. Unfortunately, there are not many methods available that explicitly calculate stationary distributions for non-complex balanced systems. Thus, most examples not fulfilling Assumption 3 are analytically intractable. However, for some calculations of stationary distributions for non-complex balanced systems, see [1]. □

5 EnvZ/OmpR signaling system

As another application of our results, we consider the two-component EnvZ/OmpR osmoregulatory signaling system in Escherichia coli, using the model proposed in [21] and considered in [3, 20]. The model corresponds to the following mass action system:

$$XD \underset{\kappa_2[D]}{\overset{\kappa_1}{\rightleftharpoons}} X \underset{\kappa_4}{\overset{\kappa_3[T]}{\rightleftharpoons}} XT \xrightarrow{\kappa_5} X_p$$
$$X_p + Y \underset{\kappa_7}{\overset{\kappa_6}{\rightleftharpoons}} X_pY \xrightarrow{\kappa_8} X + Y_p$$
$$XD + Y_p \underset{\kappa_{10}}{\overset{\kappa_9}{\rightleftharpoons}} XDY_p \xrightarrow{\kappa_{11}} XD + Y\ ,$$

where $X = \text{EnvZ}, Y = \text{OmpR}, X_p = \text{EnvZ-P}, Y_p = \text{OmpR-P}, D = \text{ADP}$, and $T = \text{ATP}$. ADP and ATP are assumed to be in large enough quantity so that their consumption in the first chain of reactions only negligibly changes their concentration. The first chain of reactions describes the phosphorylization of EnvZ, the second chain corresponds to the transfer of the phosphate group from EnvZ to OmpR, and finally the third chain describes the dephosphorylization of OmpR.

Due to [3, 20], it is known that the species Y_p exhibits ACR. Moreover, Y_p is the only ACR species. For simplicity, instead of ordering the species, here we will write the species as subscripts to refer to the corresponding entry. At equilibrium

$$z_{Y_p} = \frac{\kappa_1\kappa_3\kappa_5(\kappa_{10}+\kappa_{11})[T]}{\kappa_2(\kappa_4+\kappa_5)\kappa_9\kappa_{11}[D]} := q$$
$$z_{XT} = \frac{\kappa_3[T]}{\kappa_4+\kappa_5} z_X = \frac{\kappa_1}{\kappa_4+\kappa_5} z_{XD} = \frac{\kappa_8}{\kappa_5} z_{X_pY} = \frac{\kappa_{11}}{\kappa_5} z_{XDY_p}$$
$$z_{X_p} z_Y = \frac{\kappa_7}{\kappa_6} z_{X_pY} + \frac{\kappa_{11}}{\kappa_6} z_{XDY_p},$$

where z_S denotes the concentration of the species S. There are two quantities that are conserved at any time point, namely

$$c_1 = z_Y(t) + z_{Y_p}(t) + z_{X_pY}(t) + z_{XDY_p}(t)$$
$$c_2 = z_{X_p}(t) + z_{XT}(t) + z_X(t) + z_{XD}(t) + z_{X_pY}(t) + z_{XDY_p}(t),$$

for some positive constants c_1, c_2 depending on the initial conditions. If the amounts c_1 and c_2 are increased, then the equilibrium concentrations of all the species not exhibiting ACR are increased as well, except for X_p and Y, the equilibrium concentration of one of which could remain small.

Consider now the above reaction system in the stochastic setting. We want to know what happens if we increase the initial counts of the species such that the conserved amounts are equally increased and the initial condition is in a neighbourhood of an equilibrium point of the system. Therefore, we uniformly increase the counts of the species not exhibiting ACR, and we choose to keep X_p or Y small. We consider a sequence of processes X^N indexed by $N \in \mathbb{N}$, which are associated with the above reaction system. We assume that $X^N(0)$ is such that the entries relative to Y_p and Y, denoted by $X^N_{Y_p}(0)$ and $X^N_Y(0)$ respectively, are bounded by a constant B, and that all the other entries, if rescaled by N, converge to some positive number. In this setting, the discrete species are Y and Y_p, we have $\beta_r = 1$ for any reaction $y_r \to y_r'$ of the system and Assumptions 1 and 2 are fulfilled. For any positive vector w of continuous species concentrations, the system $\mathscr{S}_d^w$ is given by

$$0 \underset{\kappa_6 w_{X_p}}{\overset{\kappa_7 w_{X_pY} + \kappa_{11} w_{XDY_p}}{\rightleftharpoons}} Y \qquad 0 \underset{\kappa_9 w_{XD}}{\overset{\kappa_8 w_{X_pY} + \kappa_{10} w_{XDY_p}}{\rightleftharpoons}} Y_p,$$

and Assumption 3 holds thanks to Theorem 2.2. It is not difficult to check that the reaction network is irreducible, since every state of possible counts of Y and Y_p is accessible from any other state. The complex balanced equilibrium q_d^w is given by

$$q_d^w = \left(\frac{\kappa_7 w_{X_pY} + \kappa_{11} w_{XDY_p}}{\kappa_6 w_{X_p}}, \frac{\kappa_8 w_{X_pY} + \kappa_{10} w_{XDY_p}}{\kappa_9 w_{XD}} \right),$$

where the first entry refers to Y and the second one to Y_p.

The system $\mathscr{S}_c$ is given by

$$XD \underset{\kappa_2[D]}{\overset{\kappa_1}{\rightleftharpoons}} X \underset{\kappa_4}{\overset{\kappa_3[T]}{\rightleftharpoons}} XT \xrightarrow{\kappa_5} X_p$$
$$X_p \underset{\kappa_7}{\rightleftharpoons} X_pY \xrightarrow{\kappa_8} X$$
$$XD \underset{\kappa_{10}}{\rightleftharpoons} XDY_p \xrightarrow{\kappa_{11}} XD\,,$$

with

$$\lambda_{X_p \to X_pY}(w) = \kappa_7 w_{X_pY} + \kappa_{11} w_{XDY_p}$$
$$\lambda_{XD \to XDY_p}(w) = \kappa_8 w_{X_pY} + \kappa_{10} w_{XDY_p}.$$

The equilibria of the system are the positive vectors w^* that satisfy

$$w^*_{XT} = \frac{\kappa_3[T]}{\kappa_4 + \kappa_5} w^*_X = \frac{\kappa_1}{\kappa_4 + \kappa_5} w^*_{XD} = \frac{\kappa_8}{\kappa_5} w^*_{X_pY} = \frac{\kappa_{11}}{\kappa_5} w^*_{XDY_p}.$$

If $\pi_c(X^N(0))$ is such a vector w^*, or belongs to its basin of attraction, then Assumption 4 holds. Assumption 6 also holds, since the system $\widetilde{\mathscr{S}}$ corresponds to the original EnvZ/OmpR signaling system, and the unique ACR species Y_p is discrete. By making use of the fact that the original system is mass action kinetics and by Remark 3.6, it is easy to see that the remaining assumptions of Corollary 4.1 (if $\pi_c(X^N(0))$ is an equilibrium w^*) or of Corollary 4.2 (if $\pi_c(X^N(0))$ is in the basin of attraction of an equilibrium w^*) are fulfilled, and the

results can be applied. Therefore, $X^N_{Y_p}(t)$ can be approximated on average by a Poisson random variable J with mean q, in the sense of (34) or (36). The results are in accordance with the simulations in [3].

Alternatively, we could have applied the results of this paper to the signaling network by considering Y_p as the only discrete species, therefore increasing the initial counts of all other species, and by letting X^N be the process associated with

$$XD \underset{\kappa_2[D]}{\overset{\kappa_1}{\rightleftharpoons}} X \underset{\kappa_4}{\overset{\kappa_3[T]}{\rightleftharpoons}} XT \xrightarrow{\kappa_5} X_p$$
$$X_p + Y \underset{\kappa_7}{\overset{\kappa_6/N}{\rightleftharpoons}} X_pY \xrightarrow{\kappa_8} X + Y_p$$
$$XD + Y_p \underset{\kappa_{10}}{\overset{\kappa_9}{\rightleftharpoons}} XDY_p \xrightarrow{\kappa_{11}} XD + Y \ ,$$

where κ_6 has been rescaled. With this choice of rescaling, due to Remark **??**, we have that $\beta_r = 1$ for any reaction $y_r \to y'_r$ of the system. Our results can be used to draw the same conclusion as before in this different setting.

References

[1] David F. Anderson, Gheorghe Craciun, Manoj Gopalkrishnan, and Carsten Wiuf, *Lyapunov functions, stationary distributions, and non-equilibrium potential for chemical reaction networks*, arXiv:1410.4820, 2014.

[2] David F. Anderson, Gheorghe Craciun, and Thomas G. Kurtz, *Product-form stationary distributions for deficiency zero chemical reaction networks*, Bulletin of mathematical biology **72** (2010), no. 8, 1947–1970.

[3] David F. Anderson, Germán A. Enciso, and Matthew D. Johnston, *Stochastic analysis of biochemical reaction networks with absolute concentration robustness*, Journal of The Royal Society Interface **11** (2014), no. 93, 20130943.

[4] David F. Anderson and Thomas G. Kurtz, *Continuous time Markov chain models for chemical reaction networks*, Design and Analysis of Biomolecular Circuits: Engineering Approaches to Systems and Synthetic Biology (H. Koeppl et al., ed.), Springer, 2011, pp. 3–42.

[5] Karen Ball, Thomas G. Kurtz, Lea Popovic, and Greg Rempala, *Asymptotic analysis of multiscale approximations to reaction networks*, The Annals of Applied Probability **16** (2006), no. 4, 1925–1961.

[6] Franco Blanchini and Elisa Franco, *Structurally robust biological networks*, BMC systems biology **5** (2011), no. 1, 74.

[7] Péter Érdi and János Tóth, *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*, Manchester University Press, 1989.

[8] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes: characterization and convergence*, John Wiley & Sons Inc, 1986.

[9] Martin Feinberg, *Lectures on chemical reaction networks*, Delivered at Math. Res. Cent., U. Wisc.-Mad. Available for download at `http://crnt.engineering.osu.edu/LecturesOnReactionNetworks`, 1979.

[10] Martin Feinberg and Fritz Horn, *Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspaces*, Archive for Rational Mechanics and Analysis **66** (1977), no. 1, 83–97.

[11] Fritz Horn and Roy Jackson, *General mass action kinetics*, Archive for Rational Mechanics and Analysis **47** (1972), no. 2, 81–116.

[12] Hye-Won Kang and Thomas G. Kurtz, *Separation of time-scales and model reduction for stochastic reaction networks*, The Annals of Applied Probability **23** (2013), no. 2, 529–583.

[13] Hye-Won Kang, Thomas G. Kurtz, and Lea Popovic, *Central limit theorems and diffusion approximations for multiscale markov chain models*, The Annals of Applied Probability **24** (2014), no. 2, 721–759.

[14] Robert L. Karp, Mercedes Pérez Millán, Tathagata Dasgupta, Alicia Dickenstein, and Jeremy Gunawardena, *Complex-linear invariants of biochemical networks*, J. Theor. Biol. **311** (2012), 130–138.

[15] Thomas G. Kurtz, *Strong approximation theorems for density dependent Markov chains*, Stochastic Processes and Their Applications **6** (1978), no. 3, 223–240.

[16] _____, *Averaging for martingale problems and stochastic approximation*, Applied Stochastic Analysis, Springer, 1992, pp. 186–209.

[17] _____, *Equivalence of stochastic equations and martingale problems*, Stochastic Analysis 2010, Springer, 2011, pp. 113–130.

[18] Loïc Paulevé, Gheorghe Craciun, and Heinz Koeppl, *Dynamical properties of discrete reaction networks*, Journal of mathematical biology **69** (2014), no. 1, 55–72.

[19] Peter Pfaffelhuber and Lea Popovic, *Scaling limits of spatial chemical reaction networks*, arXiv:1302.0774, 2013.

[20] Guy Shinar and Martin Feinberg, *Structural sources of robustness in biochemical reaction networks*, Science **327** (2010), no. 5971, 1389–1391.

[21] Guy Shinar, Ron Milo, María Rodríguez Martínez, and Uri Alon, *Input–output robustness in simple bacterial signaling systems*, Proceedings of the National Academy of Sciences **104** (2007), no. 50, 19931–19935.

[22] Hans J. Stetter, *Numerical polynomial algebra*, Siam, 2004.

[23] Daniel W. Stroock and S.R. Srinivasa Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), vol. 3, 1972, pp. 333–359.