Scaling analysis of stochastic chemical reaction networks

Lucie Laurence

Thesis defended on December 18, 2024 for the degree of
Doctor in Mathematics, Sorbonne Université

JURY		
Dr. Marie Doumic	INRIA Saclay	Chairwoman
Dr. Carl Graham	Ecole Polytechnique	Reviewer
Pr. Bénédicte Haas	Université Sorbonne Paris Nord	Examiner
Pr. Irina Kourkova	Sorbonne Université	Examiner
Pr. Lea Popovic	Concordia University	Reviewer
Dr. Philippe Robert	INRIA Paris	PhD Advisor
Dr. Amandine Veber	Université Paris-Descartes	Examiner

Summary

Models of stochastic chemical reaction networks (CRN) represent the time-evolution of the number of molecules of different species reacting with each other. The system is composed of a finite number of species, reacting according to a finite number of chemical reactions. Though initially studied in the 70's through a determinist setting, considering the concentration of each species, the stochastic models proved to be more pertinent when one of the species is present with only a small number of molecules.

The state of the system is given by the number of molecules of each species at time t. Associating to every reaction a velocity that satisfies the law of mass action, the system is represented by a continuous time Markov jump process, with transitions given by each possible reaction, whose rates are polynomial in the state of the process.

The choice of kinetics for the reactions create two specific features for the system that will be studied in this work: the rates of reactions polynomial in the state are the source of a variety of timescales in the processes. Besides, when one of the species is in small number, the rates of the transitions can present some discontinuities. The combination of both features are the source of interesting time evolution of the processes.

In this work, we aim to study these time evolutions. Scaling methods such as averaging principles and time changes are used to study the processes.

Chapter 2 tackles two subjects: first, the Filonov's criterion, useful to show the stability of the process associated to the CRN is introduced. We then study several examples of CRN, with associated process that exhibits interesting time evolutions starting from states in the boundary of the domain.

In Chapter 3, we study the process associated to a specific CRN. From a particular type of initial state in the boundary of the domain, we show that after a scaling, the process converges to a discontinuous Markov process, of the AIMD type (addition increase, multiplicative decrease).

In Chapter 4, we study a class of CRNs for which the different processes present a hierarchy of timescales. A limit theorem for the scaled processes is shown.

The last Chapter 5 presents a technical result on CRN with only two species, showing that the process always reaches the boundary of the domain after an integrable time.

Key words. Chemical reaction networks. Stochastic processes. Markov processes. Scaling methods. Multiple timescales. Averaging principle. Time changes. Lyapunov methods.

Résumé

Les modèles stochastiques de réseaux de réaction chimiques (CRN) représentent l'évolution temporelle du nombre de molécules des espèces présentes dans le système. Ce dernier est composé d'un nombre fini d'espèces, qui réagissent entre elles via un nombre fini de réactions chimiques. Initialement étudiés dans le cadre déterministe en considérant les concentrations de chaque espèce, les modèles stochastiques s'avèrent plus pertinents quand une des espèces est présente en petit nombre.

L'état du système est donné par le nombre de molécules de chaque espèce au temps t. En associant une vitesse à chaque réaction, suivant la loi d'action de masse, le système est représenté par un processus de saut de Markov à temps continu. Les transitions du processus sont données par les réactions, et leur taux est polynomial en l'état du processus.

Le choix de cette cinétique est à l'origine des deux spécificités de ces systèmes, qui seront étudiées dans ce manuscrit: premièrement, dû aux taux polynomiaux en l'état, le processus évolue sur plusieurs échelles de temps. Par ailleurs, sur les frontières du domaines (une espèce au moins est présente dans le système en petit nombre), les vitesses de réactions peuvent présenter des discontinuités. La superposition de ces deux particularités est à l'origine des comportements intéressants de ces processus.

Dans ce manuscrit, nous étudions l'évolution temporelle de ces processus, en utilisant des outils de "scaling", tels que le principe de "stochastic averaging" et des changement de temps.

Le chapitre 2 aborde deux problématiques : premièrement, nous introduisons le critère de Filonov, utilisé pour montrer la stabilité de ces systèmes. Puis nous étudions différents exemples de CRN, dont les processus associés présentent des évolutions temporelles intéressantes.

Dans le chapitre 3, nous étudions le processus associé à un exemple de réseau chimique. Nous montrons que, partant d'un état près du bord du domaine, le processus renormalisé converge vers un processus de Markov discontinu, de type AIMD (croissance additive, décroissance multiplicative).

Dans le chapitre 4, nous étudions une classe de CRNs pour laquelle le processus associé présente une hiérarchie d'échelles de temps. Un théorème limite est montré pour le processus renormalisé.

Le dernier chapitre 5 présente un résultat technique pour une classe de réseaux à deux espèces, et montre que si le processus part du centre du domaine, il retourne au bord en un temps intégrable, quel que soit le réseau regardé.

Mots clefs. Réseaux de réactions chimiques. Processus stochastique. Processus de Markov. Méthodes de scaling. Echelles de temps multliples. Principe d'averaging. CHangements de temps.

Contents

Publications

— L. Laurence and P. Robert. "A Scaling Approach to Stochastic Chemical Reaction Networks", Arxiv: 2310.01949, September 2024.

— L. Laurence and P. Robert. "Stochastic Chemical Reaction Networks with Discontinuous Limits and AIMD processes", Arxiv: 2406.12604, June 2024.

— L. Laurence and P. Robert. "Analysis of Stochastic Chemical Reaction Networks with a Hierarchy of Timescales", Arxiv: 2408.15697, August 2024.

— L. Laurence. "On the Recurrence Properties of Stochastic Chemical Reaction Networks with Two Species", September 2024.

CHAPTER 1

Introduction

Contents

1. Chemical Reaction Networks

A model of chemical reaction network represents a system of molecules of different species, on which different reactions can occur. For example, Michaelis-Menten enzyme-substrate model is a chemical reaction network :

$$E + S \rightleftharpoons ES \rightharpoonup E + P. \tag{1.1}$$

This system has four species E, S, ES and P, and three reactions, represented by the arrows. The reaction $E + S \rightharpoonup ES$ uses a molecule of E and a molecule of S to create a molecule of ES.

Chemical reaction network theory aims to give a mathematical framework to study the time evolution of the quantity of molecules of each species.
Formally, we define a chemical reaction network (CRN) as a triple of finite sets $(\mathcal{S}, \mathcal{C}, \mathcal{R})$. The set $\mathcal{S}$ is the set of chemical species. For a model with m species, $m \geq 1$, we will usually set $\mathcal{S} = \{1, \dots, m\}$, and the species i will often be written as S_i.
$\mathcal{C}$ is a set containing linear combinations of species called *complexes*, which will represent the reactant and product of the reactions. It is a subset of $\mathbb{N}^m$, where $\mathbb{N} = \{0, 1, 2, \dots\}$: for $y = (y_i)_{1 \leq i \leq m} \in \mathcal{C}$, we will often use the notation

$$y = y_1 S_1 + \dots + y_m S_m = \sum_{i=1}^{m} y_i S_i,$$

and the notation $\emptyset$ refers to the complex associated to the null vector of $\mathbb{N}^m$. For a complex y, the integer y_i is called the *stoichiometric coefficient* of the species S_i in y.
Finally, $\mathcal{R}$ is a finite subset of $\mathcal{C}^2 \setminus \{(y, y),\ y \in \mathcal{C}\}$ that represents the *chemical reactions*. An element $(y^-, y^+) \in \mathcal{R}$ will be written $y^- \rightharpoonup y^+$. It basically corresponds to the transformation of y_i^- molecules of S_i into y_i^+ molecules of S_i, for all i.

The vocabulary of chemistry is traditionally used in the CRN theory (species, reactions, etc.) but the models do not necessarily represent real chemical systems.

The system is not necessarily closed, and reaction creating or removing molecules without a mass conservation are allowed, as *input reactions* $\emptyset \rightharpoonup y$ for $y \in \mathcal{C}$, *output reactions* $y \rightharpoonup \emptyset$, or reactions of the form $S_1 \rightharpoonup S_1 + S_2$ or $2S_1 + S_2 \rightharpoonup S_2$.
It is assumed that each species is present in at least one complex, and each complex is involved in at least one reaction.

A CRN can be also represented by an oriented graph, called the *reaction graph*, whose vertices are the complexes and whose set of directed edges is $\mathcal{R}$.

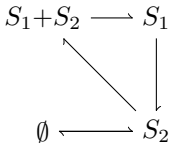


FIGURE 1. Example of a CRN

For the CRN of Figure 1, the set of species is $\mathcal{S}=\{1,2\}$, the set of complexes is $\mathcal{C}=\{\emptyset, S_1+S_2, S_1, S_2\}$ and the set of reactions is

$$\mathcal{R} = \{(\emptyset, S_2), (S_2, \emptyset), (S_2, S_1+S_2), (S_1+S_2, S_1), (S_1, S_2)\}.$$

The models of CRN can be either studied from a deterministic or a stochastic point of view, depending on the number of elements of each species present initially in the system.

When the system has a large number of molecules of each species, it is usually described by a deterministic model, which gives the time evolution of the *concentration* of each species. The state of the system is an element of the continuous space $\mathbb{R}_+^m$. These models are historically the first models of CRNs that have been studied from a mathematical point of view, they are described in Section 2.

However, when at least one of the species is present in a small number, the time evolution of the system has to be represented by a stochastic discrete model of CRN,usually a Markov process. In this setting, we study the *number of molecules* of each species, and the state of the system is an element of $\mathbb{N}^m$. These stochastic models of CRNs are the models studied in this PhD thesis. A formal definition of these models is given in Section 3.

2. Deterministic models of CRN

The deterministic model is usually used to describe systems for which all species are in large number, of the order of Avogadro's number (10^{23}) for example. The relevant quantity to represent the system is here the *concentration* of each species. The non negative real valued vector $x(t) \in \mathbb{R}_+^m$ contains the concentration of each species at time t.

To describe the evolution of $(x(t))$, one needs to define the *kinetics* for the reactions.

2.1. Mass action kinetic and dynamic. Initially introduced by the mathematician Guldberg and the chemist Waage in 1864 in [35], the law of mass action states that the velocity of a reaction is proportional to the concentration of the reactant, at the power their stoichiometric coefficient. As an example, a reaction $r_1 = S_1+S_2 \rightharpoonup S_3$ will have a velocity proportional to x_1x_2, if x_i stands for the concentration of species S_i. The positive coefficient of proportionality is noted κ_{r_1}, and is called the *constant rate* of reaction. For the reaction $r_2 = 2S_1 \rightharpoonup S_2$, the velocity will be quadratic in x_1 : of the form $\kappa_{r_2}x_1^2$.

Generally, using the notation, for $x = (x_i) \in \mathbb{R}_+^m$ and $y = (y_i) \in \mathbb{N}^m$,

$$x^y = \prod_{i=1}^{m} x_i^{y_i},$$

with the convention $0^0 = 1$, the velocity of reaction $r = y^- \rightharpoonup y^+$ in state $x \in \mathbb{R}_+^m$ is given by:

$$v_r(x) = \kappa_r x^{y^-} = \kappa_r \prod_{i=1}^{m} x_i^{y_i^-}.$$

At time t, the velocity of reaction r depends only on the state of the system. This choice of kinetics is usually chosen when the considered model is *well stirred*: the molecules are well mixed in the system, and the spatial dimension does not have to be taken into account. We will always assume the mass action kinetics for the models of CRN studied.

To a CRN, we can associate a set $(\kappa_r, r \in \mathcal{R})$, where κ_r is the constant of reaction of r. We will use the notation $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \kappa)$ to designate the CRN with mass action kinetics.

The ordinary differential equation (ODE) associated to this CRN is deduced from this kinetic: for $t \geq 0$, for all $1 \leq i \leq m$,

$$\frac{\mathrm{d}x_i}{\mathrm{d}t}(t) = \sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} v_r(x(t))(y_i^+ - y_i^-) = \sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} \kappa_r (x(t))^{y^-} (y_i^+ - y_i^-). \tag{1.2}$$

We will see later that some properties of mathematical models of CRN do not depend on the specific choice of κ_r, as long as they are positive. In fact, part of the theory aims to find such properties that depend only on the graph of the CRN (provided that the mass action kinetic is chosen). In the following, when not mentioned otherwise, we will assume that the $\kappa_r = 1$ for all reactions.

2.2. Preliminaries of CRN theory. The purpose of deterministic CRN theory is to give some insights on the solutions of ODE (1.2) : existence, uniqueness and stability of a equilibrium states are investigated.

First, note that a solution of ODE (1.2) is not necessarily defined for all $t \geq 0$. As an example, if $(x(t))$ is solution of the ODE associated to the CRN

$$2S_1 \rightharpoonup 3S_1,$$

and $x(0) = x_0 > 0$,

$$x(t) = \frac{1}{\frac{1}{x_0} - t}, \quad \text{for} \quad t \in [0, 1/x_0),$$

and $(x(t))$ goes to infinity in a finite time. The solution is only defined on a finite interval $[0, 1/x_0)$. This is one of the difficulties raised by polynomial ODEs.

The solution $(x(t))$ of the ODE (1.2) lives in a subspace of $\mathbb{R}_+^m$, called the *stoichiometric compatibility class*: for $t \geq 0$,

$$x(t) \in (x(0) + S) \cap \mathbb{R}_+^m,$$

where S is the vector space generated by $\{y^+ - y^-, y^- \rightharpoonup y^+ \in \mathcal{R}\}$:

$$S \overset{\text{def.}}{=} \left\{ \sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} a_r (y^+ - y^-),\ (a_r) \in \mathbb{R}^{\mathcal{R}} \right\}. \tag{1.3}$$

When mentioning the "uniqueness" or the "stability" of an equilibrium, we always refer to uniqueness and stability *within each stoichiometric compatibility class.*

One of the goals of the CRN theory is to find some conditions on the *graph* of the CRN to ensure the existence, uniqueness and stability of a positive equilibrium

for the ODE (1.2). A condition on the graph is by definition independent of the constants $(\kappa_r, r \in \mathcal{R})$ chosen.

A central result in this field is the Deficiency zero theorem. To state it, we need first to introduce some definitions of properties verified by the graph of a given CRN.

The connected components of the graph of a CRN determine a partition of the complexes into *linkage classes*. For example, the CRN

$$S_1 \rightleftharpoons \emptyset, \quad 2S_1 \rightleftharpoons S_2 + S_1 \rightharpoonup S_2 \tag{1.4}$$

has two linkage classes, $\{S_1, \emptyset\}$ and $\{2S_1, S_2 + S_1, S_2\}$.

A *path* from complex y to complex y' is a finite sequence of reactions that goes from y to y': it corresponds to a directed path in the graph of the CRN. For example, in the CRN (1.4), there is a path from $2S_1$ to S_2, given by $2S_1 \rightharpoonup S_2 + S_1 \rightharpoonup S_2$. A CRN is called *weakly reversible* if for $y, y' \in \mathcal{C}$, if there is a path from y to y', then we can find a path from y' to y.

The CRN of Figure 1 is weakly reversible. However, the CRN of Relation (1.4) is *not* weakly reversible, since there is a path from $2S_1$ to S_2 but no path from S_2 to $2S_1$.

The *deficiency* of the CRN is the integer

$$\delta \stackrel{\text{def.}}{=} |\mathcal{C}| - \ell - s, \tag{1.5}$$

where ℓ is the number of linkage classes of the CRN, and s is the dimension of the vector space S introduced in Relation (1.3). It can be shown (see Section 1.B.5 of the Appendix) that the deficiency is a non negative integer. For example, the CRN of Figure 1 has deficiency $\delta = 4 - 1 - 2 = 1$, whereas the CRN given in Figure 2 has deficiency 0.

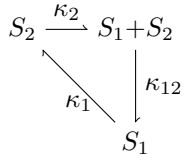


FIGURE 2. A weakly reversible CRN, with deficiency 0.

We can now state a simple version of the Deficiency zero theorem.

2.3. Deficiency zero theorem.

THEOREM 1.1. *Let $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$ be a weakly reversible CRN with deficiency zero. For any constant of reactions κ, the solution of the ODE associated to the deterministic CRN with mass action kinetic $\{\mathcal{S}, \mathcal{C}, \mathcal{R}, \kappa\}$ has precisely one equilibrium in each of its positive stoichiometric compatibility class. Besides, that equilibrium is asymptotically locally stable.*

This theorem has been established by Feinberg in [28] and [25]. Note that the full statement of Deficiency zero theorem actually says more than what is given here, see the cited reference for the original theorem.

A sketch of the proof of this theorem is given in the Appendix 1.B.

3. Stochastic models of CRN

As mentioned above, the deterministic model of CRN is well suited when *all* species are present in the system in large numbers. However, when at least one of the species of the system is present in small numbers, its time evolution is usually

better represented by a stochastic discrete model of CRN. In this configuration, we study a process $(X(t))$ on $\mathbb{N}^m$, that represents the *number of molecules* of each species at time t.

3.1. Kinetic of the reactions. Similarly to the deterministic models, we will use here the law of mass action to define the kinetics of the model. In the spirit of the heuristic reasoning of Guldberg and Waage, see Voit et al. [76], we want to find for each reaction a kinetic that would depend only on the state of the system, i.e. on the number of molecules of each species at the given time. We give here an intuitive approach to the definition of the kinetic.

If we are at state $x \in \mathbb{N}^m$, lets focus on a reaction r of the form $S_1 + S_2 \rightharpoonup y$, where y stands for any complex (different from $S_1 + S_2$). Assume that two given molecules of S_1 and S_2 meet and react after an exponential time with rate κ_r (written $\mathcal{E}(\kappa_r)$). Since there are x_1x_2 pairs of such molecules in the system, we would have such a reaction r happening after a time with exponential distribution $\mathcal{E}(\kappa_r x_1 x_2)$.

Similarly, for a reaction r_2 of the form $S_1 \rightharpoonup y$, one molecule of species S_1 would react with this reaction after a time $\mathcal{E}(\kappa_{r_2})$. Since there are x_1 molecules of S_1 in the system, we would have reaction r_2 happening after a time with distribution $\mathcal{E}(x_1\kappa_{r_2})$.

Finally, if we look at a reaction r_3 of the form $2S_1 \rightharpoonup y$, since there are only $x_1(x_1-1)/2$ such pairs, with the same ideas we will get a reaction after a time $\mathcal{E}(\kappa_{r_3} x_1(x_1-1)/2)$.

We can generalize this principle to a reaction $r = y^- \rightharpoonup y^+$: setting

$$\begin{cases} \lambda_r(x) = \kappa_r \prod_{i=1}^m \frac{x_i!}{(x_i - y_i^-)!} & \text{if for all } i, \quad x_i \geq y_i^-, \\ \lambda_r(x) = 0 & \text{otherwise,} \end{cases} \tag{1.6}$$

the reaction r will happen after a time $\mathcal{E}(\lambda_r(x))$ (the division by $y_i^-!$ is here contained in the constant κ_r). The following notation is often used :

$$x^{(y^-)} = \begin{cases} \prod_{i=1}^m \frac{x_i!}{(x_i - y_i^-)!} & \text{if for all } i, \quad x_i \geq y_i^-, \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

Note that if the state $x \in \mathbb{N}^m$, is such that $x_i < y_i^-$ for some species i, i.e. if there are not enough molecules of species i to do reaction r, $\lambda_r(x) = 0$ and the reaction can't happen. This ensures the non-negativity of the reachable states.

This allows us to define formally the process $(X(t))$. Since the rate of the reactions are only dependent on the state of the system, the process is a Markov process. It is well defined by a continuous time Markov process on $\mathbb{N}^m$, with transitions given by, for $r = y_r^- \rightharpoonup y_r^+ \in \mathcal{R}$:

$$x \mapsto x + y_r^+ - y_r^- \quad \text{at rate} \quad \lambda_r(x). \tag{1.8}$$

This kinetic is relevant as long as the system is "well stirred", since the kinetic of the reactions do not have a spatial dependance.

3.2. Specificity of the stochastic model of CRN. A process associated to a stochastic CRN with mass action kinetics exhibit several specific features. First, the rates of transitions being polynomial in the state variable, the time between two jumps of the process cannot be lower bounded in general and therefore may lead to explosion for some CRNs. As an example, the process associated to the CRN $2S_1 \rightharpoonup 3S_1$, represents a pure birth process, and starting from $x_0 \geq 2$, it explodes

almost surely in a finite time whose distribution is given by

$$\zeta = \sum_{k=x_0}^{+\infty} \frac{E_k}{k(k-1)} < +\infty$$

where (E_i) are i.i.d. random variables exponentially distributed with rate 1.

In particular, it is not clear if the process $(X(t))$ is stochastically bounded on a time interval $[0, T]$ for $T > 0$, and we don't necessarily have

$$\mathbb{P}\left(\sup_{t<T} \|X(t)\| < \infty\right) = 1.$$

Polynomial rates will also create different timescales for the different reactions. For example, in the following CRN,

$$2S_1 \rightharpoonup \emptyset \rightharpoonup S_1 \leftharpoonup S_2,$$

when the initial state is of the form (N, N) with N large, reaction $2S_1 \rightharpoonup \emptyset$ is most likely to happen before the other reactions, and it will even happen $O(N)$ times before another reaction may happen. The timescales associated to each of these three reactions are very different.

A second specificity of this choice of transition rates is seen on the boundary of the space of states. If one of the species has only a few molecules in the system, some reaction may not happen. For example, if we start from the initial state $(x_1, x_2) = (0, x_2)$ in the CRN

$$\emptyset \rightharpoonup S_1 + S_2 \rightharpoonup S_1 \rightharpoonup \emptyset,$$

the only reaction that can happen is $\emptyset \rightharpoonup S_1 + S_2$, since both other reactions need a molecule of S_1. We will see that this particularity of the stochastic models is at the origin of very interesting behaviors for the sample paths of the process.

These two features of the models of stochastic CRN with mass action kinetic are the source of difficulties, in particular to prove their stability.

3.3. Stability of a stochastic CRN. Similarly, as in the study of deterministic CRN, one of the main interest is the *stability* property of these models. Since the mathematical model is a continuous time Markov jump process, the term "stability" refers to the positive recurrence of the associated Markov process. Assuming that the process is irreducible on $\mathcal{E}_0 \subset \mathbb{N}^m$, it is positive recurrent if one can find an invariant distribution $\pi_{\mathcal{E}_0}$ on $\mathcal{E}_0$, solution of the equation, for all $x \in \mathcal{E}_0$,

$$\sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} \pi_{\mathcal{E}_0}(x + y^- - y^+)\lambda_r(x + y^- - y^+) = \pi_{\mathcal{E}_0}(x) \sum_{r \in \mathcal{R}} \lambda_r(x),$$

and, in this case, the following limit holds :

$$\lim_{t \to +\infty} \mathbb{P}\left(X(t) \in A\right) = \pi_{\mathcal{E}_0}(A), \quad \forall A \subset \mathcal{E}_0.$$

A lot of studies have been carried out on the stochastic CRN that try to give a general criterion *on the graph* of the CRN that would ensure its stability. See for example Anderson and Kim [8], Anderson et al. [9], [10].

The most famous result on this question is the (stochastic) Deficiency zero theorem, which is a consequence of the eponymous stability theorem for the deterministic CRNs (see Theorem 1.1), due to Anderson et al. [7].

The Deficiency zero theorem for stochastic models shows that if the graph of the CRN verifies the same conditions as in Theorem 1.1, i.e. if the CRN is weakly reversible and has deficiency zero, then, whatever the (positive) constant rates of the reactions are, the process associated is positive recurrent, and it even gives an explicit expression of its invariant distribution. See Anderson et al. [7].

THEOREM 1.2. *Let $(\mathcal{S},\mathcal{C},\mathcal{R})$ be a weakly reversible CRN with deficiency zero and κ a set of reaction constant. Let $(X(t))$ be the Markov process associated to the stochastic CRN with mass action kinetic $(\mathcal{S},\mathcal{C},\mathcal{R},\kappa)$, whose Q-matrix is given by (1.8), irreducible on $\mathcal{E}_0 \subset \mathbb{N}^m$. The positive measure on $\mathcal{E}_0$, defined by*

$$\pi^c_{\mathcal{E}_0}(x) = \prod_{i=1}^{m} \frac{c_i^{x_i}}{x_i!}, \quad \forall x \in \mathcal{E}_0, \tag{1.9}$$

where c is an equilibrium of the determinist *CRN with mass action $(\mathcal{S},\mathcal{C},\mathcal{R},\kappa)$, is an invariant measure of $(X(t))$.*

See Appendix 1.B for the proof.

This powerful result applies only to a class of specific CRN, and outside of this class, almost no stability results stands, though the following conjecture is believed to be true :

CONJECTURE 1.3. *A continuous time Markov jump process associated to a* weakly reversible *stochastic CRN with mass action kinetic is always positive recurrent.*

It has not been proven until now.

However, several weaker results have been shown, and tools have been developed to try to show this conjecture, all held on the Lyapunov theorem.

4. Mathematical tools

This Section is dedicated to the presentation of several mathematical tools used to study stochastic CRNs.

In the following, unless otherwise specified, $(X(t))$ is a continuous time Markov process on $\mathbb{N}^m$ associated to a stochastic CRN with mass action kinetic $(\mathcal{S},\mathcal{C},\mathcal{R},\kappa)$, with $\mathcal{S} = \{1,\dots,m\}$.

4.1. Stability criterion. To show the stability of the process associated to a stochastic CRN, and if the Deficiency zero theorem does not hold, the ideas of a Lyapunov criterion can be useful. It consists in showing that the "energy" of the system is decreasing: Here, we will call *Energy function* a non negative function on the set of states $\mathbb{N}^m$, that goes to infinity when the state gets large, i.e. that verifies

$$\lim_{\|x\|\to+\infty} f(x) = +\infty.$$

The Energy function used the most in the CRN theory community is the Entropy function, defined as

$$V(x) = \sum_{i=1}^{m} x_i \ln(x_i) - x_i + 1.$$

We will also use as energy function the norm L^1, $\|\cdot\|_1$, linear functions of the form $x \mapsto a_1x_1 + \ldots + a_mx_m$, with $a_i > 0$ for all i, or polynomials of the form $x \mapsto a_1x_1^{q_1} + \ldots + a_mx_m^{q_m}$ for $q_i \geq 1$.

The goal is to define the energy of the system such that it decreases with time, if the initial state is large enough. This method is often used in the literature of stochastic CRN, with the Foster-Lyapunov criterion. Let $\mathcal{A}$ be the infinitesimal generator of the process with Q-matrix given by (1.8) :

$$\mathcal{A}(f)(x) = \sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} \lambda_r(x)(f(x+y^+-y^-) - f(x)), \tag{1.10}$$

for any function $f : \mathbb{N}^m \to \mathbb{R}$ and $x \in \mathbb{N}^m$.

The Foster-Lyapunov criterion is the following : if one can find an energy function f_0, a finite subset $K \subset \mathbb{N}^m$ and some $\gamma > 0$ such that

$$\mathcal{A}(f_0)(x) \leq -\gamma \quad \text{for} \quad x \in \mathbb{N}^m \setminus K,$$

then the process is positive recurrent.

This criterion has been extensively used in Anderson and Kim [8], Anderson et al. [9], Anderson et al. [10] to show the stability of some classes of stochastic CRN, using the Entropy function as an Energy function.

4.2. Scaling analysis.

4.2.1. *Motivation.* Knowing that a process is stable, or even the explicit expression of its invariant distribution only gives information on its the long term behavior, and not on its transient behavior. If we are looking for some results on the sample paths of the process, or the value of a hitting time (time needed by the process to reach a set $A \subset \mathbb{N}^m$ from a state x), some other methods have to be used.

One of the tools to get this kind of information is a scaling approach. The idea is to take some "scaling parameter" N and study the process $(X^N(t))$ depending on N. The parameter N can be the norm of the initial state of $(X^N(t))$, or the volume of the system... We are looking for some limit results on functions of $(X^N(t))$ when N gets large.

To find some limit theorems on the sample paths of the process, we will need to "scale" the process according to the parameter N, in time and/or in space.

Usually, for CRNs, the *scaled process* will take the form

$$\left(\overline{X}^N(t)\right) = \left(\frac{X_i^N(t/N^\gamma)}{N^{\alpha_i}},\ 1 \leq i \leq m\right),$$

where $\gamma, \alpha_i \in \mathbb{R}$. The scaling in space can be different for each species. For a species i, α_i is chosen sufficiently large to ensure that the process $(\overline{X}_i(t))$ stays bounded, but not too large so that this process does not become trivial when N gets large. The scaling in time may depend on the regime one wants to investigate since different timescales can be of interest for the same model.

4.2.2. *Stochastic differential equations (SDE).* The continuous time Markov process given by the transitions (1.8) can be expressed as the solution of a Stochastic Differential Equation (SDE). This representation is especially useful to establish functional limit theorems. See Theorem (20.6) of Rogers and Williams [69].

For a Poisson process $\mathcal{P}$ on $\mathbb{R}_+^2$, with intensity measure the Lebesgue measure on $\mathbb{R}_+^2$, we will use the notation for $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\mathcal{P}(A, \mathrm{d}t) = \int_{x \in A} \mathcal{P}(x, \mathrm{d}t).$$

The use of Poisson process on $\mathbb{R}_+^2$, instead of on $\mathbb{R}$ to write the SDE is particularly useful when the process has time dependent, random intensities. As an example, the process $(X(t))$ that represents the CRN $S \rightharpoonup 2S$ verify the Relation

$$X(t) = X(0) + Y\left(\int_0^t X(s)\,\mathrm{d}s\right),$$

where Y is a Poisson process on $\mathbb{R}$. This is Kurtz's representation, see Kurtz [51] Using a Poisson process $\mathcal{P}$ on a 2 dimensional space, the previous equation is equivalent to

$$X(t) = X(0) + \int_0^t \mathcal{P}((0, X(S-)), \mathrm{d}s).$$

For a general CRN, the SDE verified by the process $(X(t))$ whose transitions are (1.8) is :

$$\mathrm{d}X(t) = \sum_{r=(y_r^-,y_r^+)\in\mathcal{R}} \left(y_r^+ - y_r^-\right) \mathcal{P}_r\left(\left(0, \kappa_r \frac{X(t-)!}{(X(t-)-y_r^-)!}\right), \mathrm{d}t\right),$$

where $(\mathcal{P}_r, r \in \mathcal{R})$ are independent Poisson processes on $\mathbb{R}_+^2$ with intensity measure the Lebesgue measure on $\mathbb{R}_+^2$.

4.2.3. *Classical scaling.* This historical example was meant to show some connection between the deterministic and the stochastic model of CRN. The stochastic point of view is usually used when at least one of the species has a small number of molecules. Therefore, when *all* species are in large number, we could expect the stochastic model to "converge" in some way to the determinist model, solution of an ODE.

This can be formulated with a scaling according to *the volume* of the system V. It has initially been presented in Kurtz [51]. Each species of the system is present initially with a number of molecules of order V: $X_i^V(0) = O(V)$. The scaling in space is the same for all species:

$$\overline{X}^V(t) = \frac{X^V(t)}{V}$$

The particularity of this system is that the reaction constants are also scaled with the volume. Defining the reaction rates as in Section 3.1, we understand them as a "frequency of meeting" for the molecules needed for the reaction. If the volume of the system increases, the frequency of meeting should decrease. More precisely, the more molecules are needed for the reaction, the more it will be hard for them to meet. Conversely, for the input reactions, if the volume of the system increases, there are more ways to enter the system, and therefore the frequency of the reactions $\emptyset \rightharpoonup y$, for $y \in \mathcal{C}$, increases. We define the scaled constant of reactions as follows: for a reaction $r = y^- \rightharpoonup y^+ \in \mathcal{R}$,

$$\bar{\kappa}_r^V = \frac{\kappa_r}{V^{\|y^-\|-1}},$$

where $\|y\| \overset{\text{def.}}{=} \sum_{i=1}^m |y_i|$ for $y \in \mathbb{R}^m$.

Note that if the reaction needs only one molecule to happen, $\|y^-\| = 1$ and no scaling is done.

This scaling done, all reaction rates are of order of V. Indeed, if the process is at a state $x_V = (x_1 V, \ldots, x_m V)$, for a given reaction $r = y^- \rightharpoonup y^+$,

$$\lambda_r^V(x_V) = \bar{\kappa}_r^V x_V^{(y^-)} \approx \frac{\kappa_r}{V^{\|y^-\|}} \prod_{i=1}^m (x_i V)^{y_i^-} = \kappa_r V x^{y^-}.$$

Since the scaled process needs a number of reactions of order V to change significantly, and every reaction happens at a frequency of order V, the normal timescale $t \mapsto t$ is the right timescale to study the process $(\overline{X}^V(t))$. Here, no scaling in time is necessary.

In this configuration, it can be shown that the scaled process $(\overline{X}^V(t))$ converges to a deterministic function, solution of the ODE associated to the deterministic CRN, with mass action kinetic and reaction constant (κ_r), stated in (1.2). This result, known as the "classical scaling" of CRNs, has been formally proved in Kurtz [51]. See also Section 4.1 of Chapter 2.

4.2.4. *Multiple timescales and averaging principle.* In the previous example, the scaling on the κ_r and the "uniform" scaling in space put us in a very convenient situation, where the processes of each species evolves at the *same timescale*. However, in more general scaling situations, this is not always the case and one single scaling will not always allow us to investigate the behavior of each species.

For example, lets look at the CRN of Figure 2. The process $(X(t))$ associated to this CRN has the following transitions :

$$x \in \mathbb{N}^2 \to \begin{cases} x - e_2 & \text{with rate} \quad \kappa_{12} x_1 x_2, \\ x + e_2 - e_1 & \text{with rate} \quad \kappa_1 x_1, \\ x - e_1 & \text{with rate} \quad \kappa_2 x_2, \end{cases}$$

Starting from a state $x^N = (N, N)$ with $N \gg 1$, we would like to see the evolution of the process scaled in space

$$\left(\frac{X_1^N(t)}{N}, \frac{X_2^N(t)}{N} \right).$$

(a) **Decrease of** $(X_2^N(t)/N)$**:**The process $(X_2^N(t)/N)$ has a transition $x_2 \to x_2 - 1/N$ with rate $O(N^2)$, associated to the reaction $S_1 + S_2 \rightharpoonup S_1$, therefore it evolves at the timescale $t \mapsto t/N$. On the other hand, since both reactions changing $(X_1^N(t))$ have a rate $O(N)$, at this timescale, the process $(X_1^N(t/N)/N)$ *does not change*. Here, $(X_1^N(t)/N)$ is a *slower process* than $(X_2^N(t)/N)$. The process $(X_1^N(t/N)/N)$ converges to a *constant* equal to 1 when N gets large, and this result leads to the convergence of $(X_2^N(t/N)/N)$, to the function $t \to e^{-\kappa_{12} t}$. Here the multiplicity of timescales is not an issue, since both processes are "well known" at the timescale we chose: one moves slowly, and the other is constant.

(b) **Decrease of** $(X_1^N(t)/N)$**:** Once $(X_2^N(t))$ has decreased, we can take the initial state of the process as (N, k), for some $k \in \mathbb{N}$, and look at the process scaled in space $(X_1^N(t)/N, X_2^N(t))$. The process $(X_2^N(t))$ has transitions with rate $O(N)$ and therefore moves on a timescale $t \mapsto t/N$. On the other hand, the process $(X_1^N(t)/N)$ has transitions $x \mapsto x - 1/N$ at rate N, and therefore it evolves on the normal timescale $t \mapsto t$. To have some scaling results on $(X_1^N(t)/N)$, one has to deal with $(X_2^N(t))$ which is here a *fast process*. In this situation, we use an *averaging principle*. The slow process is expressed in terms of the equilibrium of the fast process. A convergence of the fast process is obtained for its *occupation measure*. Here, the occupation measure (μ_N) associated to $(X_2^N(t))$ is given by

$$\langle \mu_N, g \rangle = \int g(s, X_2^N(s)) \, \mathrm{d}s.$$

The averaging principle consists in proving the convergence in distribution of the pair "slow process + occupation measure of the fast process" $((X_1^N(t)/N), \mu_N)$, and in finding an expression for its limit. Here, the limit of $(X_1^N(t)/N)$ is $(e^{-\kappa_1 t})$ and the limit of μ_N is the measure $\mathrm{d}s \otimes \pi(x)$ where π is a Poisson distribution on $\mathbb{N}$ with parameter κ_1/κ_{12}.

The limit of the slow process is sometimes expressed as the solution of an ODE involving the limit of the occupation measure. The proof of the convergences of this example can be found in Section 6 of Chapter 2.

We review quickly the literature in a stochastic context. It should be noted that this is also a well-studied topic in (deterministic) dynamical systems. See Verhulst [75] for example. Early works on the proof of averaging principles are due to Has'minskiĭ. See Has'minskiĭ [39, 40]. Chapter 7 of Freidlin and Wentzell [31]

considers these questions in terms of the convergence of Cesaro averages of the fast component. Papanicolaou et al. [63] has introduced a stochastic calculus approach to these problems, mainly for diffusion processes. Kurtz [55] has extended this approach to jump processes by giving convenient (and useful) tightness criteria for sequences of random measures and a characterization of their possible limiting points.

Averaging principles have been used in various situations to study chemical reaction networks. Kurtz and co-authors have in particular investigated several examples of CRNs with scaling methods. In this approach, some reaction rates may be sped-up with some power of the scaling parameter and the state variables are scaled accordingly. See for example Ball et al. [11], Kang and Kurtz [46], and Kim et al. [50] where, for various examples of CRNs, the choice of convenient scalings of reaction rates is investigated and limit theorems are derived. The technical analyses of these references follows the general lines of Kurtz [55].

4.2.5. *Random time change.* One of the specific features of stochastic CRNs is the discontinuity of the kinetics on the boundary of the space of states, which can create discontinuities in the timescale of the process. As an example, lets look at the following CRN :

$$\emptyset \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} S_1 \overset{\kappa_3}{\leftharpoonup} S_1 + S_2.$$

Starting from the state $(1, N)$, with $N \gg 1$, as long as $X_1^N(t) \geq 1$, $X_2^N(t)/N$ decreases exponentially at the normal timescale. However, if reaction $S_1 \rightharpoonup \emptyset$ decreases $X_1^N(t)$ to 0, one will have to wait for reaction $\emptyset \rightharpoonup S_1$, i.e. for a time exponentially distributed to see any new evolutions on $X_2^N(t)$. $(X_2^N(t)/N)$ is alternatively constant, and exponentially decreasing.

This discontinuity can create difficulties to show any limit theorem for the sample path of the scaled process. One way of dealing with it is to use random time changes, to separate the two different evolutions. In this example, we can simply "cut out" the times where $X_1^N(t) = 0$, which does not take any information away, do the scaling analysis on the process $(Y^N(t))$ that we get once these delays are taken off, and finally add the delays a posteriori to reconstruct the process.

Formally, we define a time change $(\tau^N(t))$, and a process $(Y^N(t)) \overset{\text{def.}}{=} (X^N(\tau^N(t)))$ on which the usual techniques of convergence in distribution can be applied. The idea is to show the convergence in distribution of $(\tau^N(t))$ to $(\tau(t))$, and a scaled version of $(Y^N(t))$ to $(y(t))$. Under some convenient conditions, one can then deduce the convergence of the scaled version of $(X^N(t))$ to $(y(\tau^{-1}(t)))$, where $(\tau^{-1}(t))$ stands for the generalized inverse of $(\tau(t))$.

In the example presented above, we set

$$\tau^N(t) = \int_0^t \mathbb{1}_{\{X_1^N(s) \geq 1\}} \, \mathrm{d}s.$$

We get a convergence in distribution of the process $(Y_2^N(t)/N) = (X_2^N(\tau^N(t))/N)$ to $(e^{-t\kappa_3})$, and the limit $(\tau(t))$ is here a discontinuous Markov process.

For a technical framework of random changes in time, see Chapter 6 of Ethier and Kurtz [24] and Section III.21 of Rogers and Williams [68]. A similar example is studied in detail in Chapter 2

4.3. Topologies on the Skorokhod space. When looking at the convergence in distribution of processes, the diversity of phenomena present in the stochastic models of CRN created some difficulties to choose the right *topology* on the space of sample paths. Indeed, if the scaled process converges, sometimes, for the uniform norm (uniform convergence) to the solution of an ODE, as in the classical scaling of Section 4.2.3, we will see later that the combination of boundary effects

and polynomial rates can create some discontinuities in the limit of the scaled process, which prevent uniform convergence of the sequences.
Here, the continuous time Markov jump processes live in the Skorokhod space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}_+^m$ (continuous from the right, with left limits at all t), on which several topologies can be defined.

Four topologies are useful in the study of stochastic CRNs: the uniform topology, the J_1 and M_1-Skorokhod topologies and the S-Jakubowski topology. A detailed description of each of the topologies is given in the Appendix 1.A, I will give here a small survey of the reason we need to weaken the uniform topology.

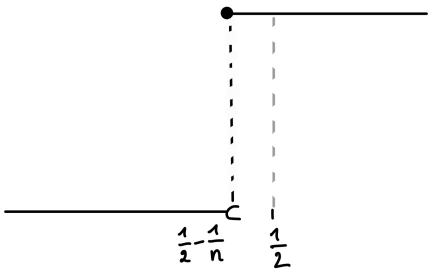


FIGURE 3. Converges in the J_1 topology

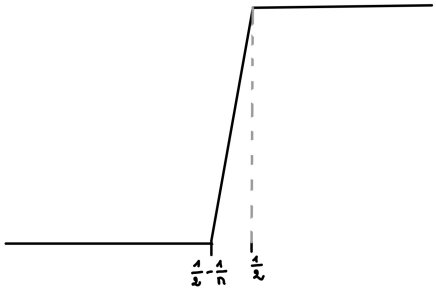


FIGURE 4. Converges in the M_1 topology

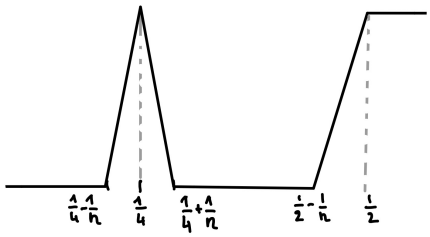


FIGURE 5. Converges in the S topology

The uniform topology works well when the limit of the process is continuous, and is most often the topology used in practice. When the possible limit of the sequence exhibits a jump, this topology is not weak enough. For example, the sequence of Figure 3 does not converge for the Uniform topology. To cope with this, Skorokhod introduced the J_1 topology, that adds some flexibility on the time instants of jumps, and for which the sequence of Figure 3 converges, to the function $\mathbb{1}_{\{[1/2,1)\}}$. This topology is usually sufficient to show the convergence of processes to a process with jumps.

In the framework of stochastic CRNs however, a final jump of a scaled process can be the consequence of a large number of small jumps happening in a small time interval, whose length goes to zero when the parameter of scaling gets large. This situation is not well apprehended by the J_1-Skorokhod topology. Indeed, for this topology, if $x_n \to x$ and x has a jump at time t, the functions x_n must have a jump of similar height at a time $t_n \to t$. As an illustration, the sequence of Figure 4 does not converge to $\mathbb{1}_{\{[1/2,1)\}}$ for the J_1-Skorokhod topology, since the x_n

are continuous, and therefore have no jump around $t = 1/2$. A weaker topology can be defined to cope with this problem : the M_1-topology, for which this particular example converges. This topology is constructed using the "extended graph" of the function, defined for a function x as the set containing the graph $\{(t, x(t))\}$ and all the points between $(t, x(t-))$ and $(t, x(t))$. The value of the functions x_n does not really mater around a instant of jump t of the limit function, as long as it is in the interval $[x(t-), x(t)]$.

Finally, one last situation can be seen in the stochastic CRNs: the graph of a scaled process can present a "spike", whose length goes to zero when the scaling parameter goes to infinity. As an example, see the sequence of Figure 5. The spike at $t = 1/4$ should intuitively converge to a continuous point with value 0, and the sequence should converge to the function $\mathbb{1}_{\{[1/2,1)\}}$. This is not the case with the M_1-Skorokhod topology, since around $1/4$ in Figure 5, the graph of x_n always have points very far from the extended graph of $\mathbb{1}_{\{[1/2,1)\}}$. To study such a convergence, one can use a weaker topology, the S-Jakubowski topology.

This last topology is a sequential topology, that cannot be metricized. However, its criterion for compactness is very simple to use, and the usual properties needed to study convergence of processes hold for this topology.

From the strongest to the weakest to the weakest, the topologies can be ordered as follow :

$$\text{Uniform} \ \rightarrow J_1 \ \rightarrow \ M_1 \ \rightarrow \ S.$$

A sequence converging for the J_1 topology converges for the M_1 and the S-Jakubowski topologies for example, but not necessarily for the uniform topology.

The advantages and drawbacks of each topology are more extensively presented in Appendix 1.A, as well as their use in the framework of convergence of processes.

For references on the uniform topology and the J_1-topology, see Billingsley [14]. For the M_1-Skorokhod topology, see Whitt [77], and for the S-Jakubowski topology, see Jakubowski [44].

5. Contributions

Two specific features characterize Markov processes associated to a stochastic chemical reaction network.

(a) **Polynomial rates :** the transition rates exhibit a polynomial dependance on the state variable, and therefore can have different orders of magnitude. This is reflected in the behavior of the process as a *multiplicity of timescales* and/or a multiplicity of order of magnitude of each species. As an example, in the CRN introduced in Agazzi and Mattingly [4],

$$S_2 \rightharpoonup \emptyset \rightharpoonup S_1 + S_2, \quad 5S_1 + 2S_2 \rightharpoonup 3S_2 \rightharpoonup 2S_1, \tag{1.11}$$

if both species are $O(N)$ with $N \gg 1$, reaction $5S_1 + 2S_2 \rightharpoonup 3S_2$ is by far the fastest reaction, since it has a rate of order N^7, and it is the only reaction to happen until $X_1^N(t)$ has decreased enough. This creates multiple relevant timescales for the process, as can be seen in the representation of the sample paths of the processes :

Figure 6 shows on the left the timescale where $(X_2^N(t)/N)$ decreases, and on the right a zoom in at time $t = 0$ to see the decrease of $(X_1^N(t)/N)$, which happens at a much faster timescale. These multiple timescales can be the source of difficulties when one wants to prove a scaling result on these processes, see the example described in Section 4.2.4.

(b) **Boundary effects :** the dynamical behavior along some boundaries of the space of states can present discontinuities, referred to as "boundary

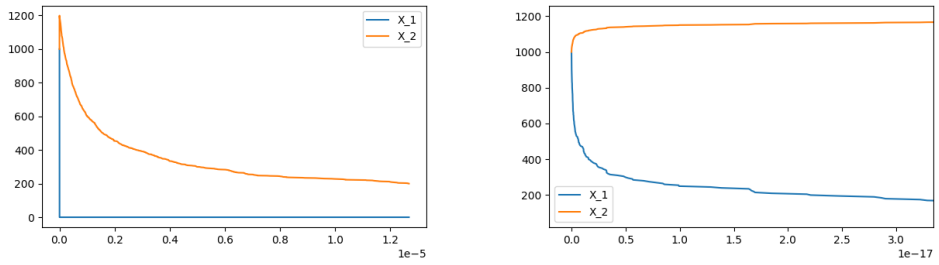


FIGURE 6. Process $(X_1^N(t), X_2^N(t))$ associated to the CRN (1.11), starting from (N, N), $N = 1000$.

effects". This is due to the fact that some chemical reactions require a minimal number of copies of some chemical species to take place. Boundary effects appear when the complexes of the CRN contain different species. For example, the CRN

$$\emptyset \rightleftharpoons S_1 \leftharpoonup S_1 + S_2.$$

have some boundary effects because of the complex S_1+S_2. Indeed, starting from the state $(0, N)$ with $N \gg 1$, even if the rate of this reaction is "globally" of the order N, in fact it depends closely on the states visited by $(X_1^N(t))$: if $X_1^N(t) = 0$, the reaction is "frozen" and can't happen, and when $X_1^N(t) \geq 1$, it has a very large rate and is the fastest reaction.

It is usually a combination of both features that creates the most interesting behaviors. The work of my PhD thesis aim to study these phenomena using scaling methods.

As presented in Section 4.2, scaling methods have already been used to study stochastic CRNs. Until now however, most of the studies, motivated by biological models, used scalings of the constant of reactions. See Kang and Kurtz [46] and Ball et al. [11] for example.

Our purpose is to study the process without changing the order of magnitude of the constant of reactions, given by the mass action kinetic. We study two types of scalings :

(a) In Chapters 2 and 3, we focus on the boundary effects of the system. The scaling parameter is chosen as the *norm of the initial state* of the process. We study interesting examples of CRN to present these boundary effects, and undertake a scaling analysis from a large state in the boundary of the system. We investigate the time evolution of the CRN, in particular how the process starting from such a saturated initial state returns to the neighborhood of the origin. See Section 5.1 for more details on the motivations. If N is the scaling parameter, the initial state x_N verifies

$$\lim_{N \to +\infty} \frac{\|x_N\|}{N} = 1,$$

where $\| - \|$ stands for the norm L^1 in $\mathbb{N}^m$, and for at least one species,

$$\lim_{N \to +\infty} \frac{x_i^N}{N} = 0.$$

The phenomena observed using this scaling on different specific CRNs proved to be diverse, and their study with scaling methods quite challenging.

(b) In Chapter 4, we study a CRN with *fast input* : the constant of input reactions (of the form $\emptyset \to y$, for a complex y) are chosen as $\kappa_r N$, where N is the parameter of scaling. All the other constants of reactions are unchanged. We study a CRN chosen to have *no boundary effects* in fact, by taking only complexes of the form kS_i, with $k \geq 1$, and focus on the multiple timescales present in the system, due to the polynomial rates. For these CRNs, the process studied lives in the interior of the state space, boundary effects do not play a role.

5.1. Chapter 2: A Scaling Approach to Stochastic Chemical Reaction Networks. In this first Chapter, two main subjects are developed.

First, we introduce an extended criterion to prove the stability of stochastic CRNs: the Filonov criterion. Second, we introduce a scaling method with respect to the norm of the initial state on several examples.

The Filonov criterion for stability. Recall that the Foster-Lyapunov criterion is often used to show the stability of a Markov jump process. It relies on the definition of an Energy function f_0, i.e. a non negative function on $\mathbb{N}^m$ that verify

$$\lim_{\|x\| \to +\infty} f_0(x) = +\infty,$$

and the proof of an inequality :

$$\mathcal{A}(f_0)(x) \leq -\gamma, \quad \text{for} \quad x \in \mathbb{N}^m \setminus K,$$

where K is a finite set of $\mathbb{N}^m$, $\gamma > 0$ and $\mathcal{A}$ is the infinitesimal generator of the process. See Section 4.1 for more details.

However, this criterion can be in practice very difficult to use. Indeed, for the Foster-Lyapunov criterion to hold, the system has to see its energy decrease after *one* jump of the process. Working with one of the "simple" energy function presented in Section 4.1, it is not always the case. For example, the CRN

$$\emptyset \rightharpoonup S_1 + S_2 \rightharpoonup S_1 \rightharpoonup \emptyset \tag{1.12}$$

is stable since it verifies the conditions of the deficiency zero theorem, but for any of the energy functions presented above, one has $\mathcal{A}(f_0)(0, x_2) \geq 0$. This is due to some boundary effects of the system: from a state $(0, x_2)$ with $x_2 \geq 1$, the *only* reaction that can happen is $\emptyset \rightharpoonup S_1 + S_2$, that increases the energy, f_0 being the entropy, a norm, or polynomial.

A first approach here would be to define the energy function piecewise on a partition of the state space. This is the approach used in Agazzi and Mattingly [4], and it raises a lot of difficulties once one needs to "glue" the different parts together.

Another way to cope with this is to use an extended version of the Foster-Lyapunov criterion, known as *the Filonov criterion*, see Filonov [30]. This is the approach that is presented in the first article. For this criterion, instead of requiring that the energy of the system decreases after the first jump of the process $(X(t))$, we allow a *delay* before the decrease. The counterpart is that the decrease must be in some way proportional to the "waiting time". The criterion is expressed as follows :

$$\mathbb{E}_x \left[f_0(X(\tau)) - f(x) \right] \leq -\gamma \mathbb{E}_x[\tau], \quad \forall x \notin K,$$

where as before, f_0 is an energy function, $\gamma > 0$ and K a finite set of $\mathbb{N}^m$, and τ is an *integrable stopping time*.

For example, on the CRN (1.12), starting from $(0, x_2)$ with $x_2 \gg 1$, the first jump is to the state $(1, x_2 + 1)$, but then the most likely reaction to happen is $S_1 + S_2 \rightharpoonup S_1$, and it happens several times in a row, since its rate is proportional to x_2. Therefore, after the third jump, the process is most likely at the state

$(1, x_2 - 1)$. Taking as an energy function $f(x) = x_1 + 2x_2$, and τ the instant of the third jump of the process, we have

$$\mathbb{E}_{(0,x_2)}[X_1(\tau) + 2X_2(\tau) - 2x_2] \leq -\mathbb{E}_{(0,x_2)}[\tau]$$

for x_2 large enough. This deals with the states at the boundary of the state space. For the other states, one can define τ as the first instant of jump of the process, and show that the inequality holds.

Note that the stopping time τ is random. It can be chosen as here as the instant of the n^{th} jump of the process, but also as the hitting time of some subset of $\mathbb{N}^m$, and can even be deterministic.

This criterion is presented in Section 4 of Chapter 2, and is applied to different examples in the four following sections. In particular, in Section 7 it is applied to a specific example of the literature on which it is particularly efficient.

To use this criterion, it is pretty straightforward to see that one has to "know" the behavior of the system, and more precisely the time evolution of its energy, to be able to set the "right" stopping time. The key information needed to apply Filonov's criterion is to understand how the energy decreases, starting from a given large state. Because of the definition of the energy function, this is closely related to the behavior of the process itself, starting from a large state. This encourages the study of the asymptotic behavior of the sample paths of the system with a scaling analysis with respect to the norm of the initial state, presented above.

Scaling analysis. This chapter presents also several scaling results, that illustrate the diversity of boundary effects on the sample paths of CRNs, and of timescales within a single CRN.

One scaling is done for a general class of CRN. It states that if all species are of order N, with N large, slowing down the time so that *only the fastest reactions* can be seen, the scaled process converges in distribution to a deterministic function solution of an ODE. This result is quite similar to the "classical scaling" presented in section 4.2.3, but without any scaling on the κ_r.

We then present specific examples of CRNs that, from our point of view, exhibit interesting behaviors.

Three scaling results are presented for the CRN of Figure 2. On this first example, we show that depending on the initial state chosen for the process, already three different timescales are relevant. Starting from (N, N) with N large, the relevant timescale is $t \to t/N$. This is in fact a consequence of the general result presented above. Starting from $(\sqrt{N}, N)$, the interesting timescale to see the evolution of $(X_2^N(t)/N)$ is here $t \to t/\sqrt{N}$. Finally, starting from $(N, 0)$, using an averaging principle, we show that the relevant timescale is the normal timescale $t \to t$.

A part is dedicated to the CRN of Agazzi and Mattingly [4] introduced in Relation (1.11), and a scaling analysis is done to better understand the sample path of the process starting from $(N, 1)$, to a neighborhood of zero.

Finally, a thorough study of the CRN

$$\emptyset \underset{\kappa_1}{\overset{\kappa_0}{\rightleftharpoons}} S_1+S_2, \qquad pS_1+S_2 \underset{\kappa_3}{\overset{\kappa_2}{\rightleftharpoons}} pS_1+2S_2,$$

with $p \geq 2$ is carried out. It has been introduced and discussed in Agazzi et al. [2] for $p=2$. Its transient behavior exhibits several unusual scaling properties with respect to the norm of the initial state. Such a behavior is probably not an exception in the context of CRNs, but more likely quite common. Additionally, it gives an interesting example of the use of time change arguments to derive scaling results.

The associated Markov process $(X_1(t), X_2(t))$ has several boundary behaviors: at least p copies of S_1 are required for the second reaction to occur, and negative

jumps of $(X_1(t))$ can occur only when $(X_2(t))$ is not null. It leads in fact to a kind of bi-modal behavior for the CRN.

A scaling analysis of this CRN is achieved for two classes of initial states.

(a) An initial state of the form (N, b), with $b \in \mathbb{N}$ fixed.
Our main result shows that there exists some $t_\infty > 0$, such that the convergence in distribution of processes

$$\lim_{N \to +\infty} \left(\frac{X_1(Nt)}{N}, t < t_\infty \right) = (C(t_\infty - t), t < t_\infty),$$

holds for some constant $C > 0$.

(b) If the initial state is of the form (a, N), $a < p$.
The occupation measure of $(X_2(N^{p-1}t)/N)$ converges to the occupation measure of $(V(t))$, where $(V(t))$ is an *explosive* Markov jump process on $(0, 1]$, with a multiplicative structure, almost surely converging to 0.

Contrary to the first examples, to decrease the norm of the process, one has to speed-up the timescale by a factor N in (a) and N^{p-1} in (b) and the decay in (a) is only linear with respect to time. This is somewhat unusual in the current mathematical literature of CRNs, but these kind of examples are more and more studied, since they are source of difficulties showing up in the study of the stability of general CRN, as mentioned in Agazzi et al. [2].

5.2. Chapter 3: Stochastic Chemical Reaction Networks with Discontinuous Limits and AIMD processes. In this Chapter, we study a class of cyclic CRNs, with m species, $m \geq 2$:

$$\begin{aligned} \emptyset \xrightarrow{\kappa_0} S_1 \xrightarrow{\kappa_1} S_1 + S_2 \xrightarrow{\kappa_2} \cdots \\ \xrightarrow{\kappa_i} S_i + S_{i+1} \xrightarrow{\kappa_{i+1}} S_{i+1} + S_{i+2} \xrightarrow{\kappa_{i+2}} \cdots \\ \xrightarrow{\kappa_{m-1}} S_{m-1} + S_m \xrightarrow{\kappa_m} S_m \xrightarrow{\kappa_{m+1}} \emptyset. \end{aligned} \tag{1.13}$$

Once again, the CRN presents boundary effects, since the reaction in the middle of Relation (1.13) does not occur if either x_i or x_{i+1} is null. A molecule of S_i may be transformed into a molecule of S_{i+2} only if there is at least a molecule of S_{i+1}, even if the $(i+1)$th coordinates is not changed by the reaction.

We investigate this class of CRNs via the convergence in distribution of its scaled sample paths. In the spirit of Chapter 2, the scaling we consider in this chapter does not change the basic dynamic of the CRN, in particular its reaction rates. It is assumed that the initial state of the CRN is "large", its norm is proportional to some scaling parameter N. We investigate the time evolution of the process associated to this CRN, starting from a saturated initial state and its way back to a neighborhood of the origin.

A first Section is dedicated to the study of the CRN with a general m. Because of the polynomial rate, if both S_i and S_{i+1} are of order N, the reaction $S_i + S_{i+1} \rightharpoonup S_{i+1} + S_{i+2}$ has a rate of order N^2, which is maximal for this class of CRNs. We show that the process goes "quickly" to a set of states for which the indices of the coordinates whose value is greater than εN are at distance at least two, the other coordinates being $o(N)$.

In a second Section we study the CRN with three chemical species. Several interesting initial states are studied, and for each one, a scaling analysis is carried out. In particular, starting from a state $x_N = (1, N, 1)$, it is shown that the process $(X_N(t)/N)$ converges in distribution to a continuous, but random, process. This case provides an example of a CRN whose first order is not the solution of a set of deterministic ODEs.

The three last parts are dedicated to the study of the CRN with four chemical species. A class of initial states gives rise to a more complex behavior than what we have observed when m=3. We did not try a complete (cumbersome) classification of initial states here, but we do believe that the class of initial states studied exhibits the most interesting behaviors.

$$\emptyset \xrightarrow{\kappa_0} S_1 \xrightarrow{\kappa_1} S_1+S_2 \xrightarrow{\kappa_2} S_2+S_3 \xrightarrow{\kappa_3} S_3+S_4 \xrightarrow{\kappa_4} S_4 \xrightarrow{\kappa_5} \emptyset.$$

The initial states considered are of the type $(0,N,0,0)$, with the convention that "0", resp. "N", means $O(1)$, resp. $O(N)$. We show that the process lives in the subset of the state space of elements of the type $(0,N,0,\sqrt{N})$ and that the decay of the norm of the state occurs on the timescale $(\sqrt{N}t)$. More important, this decay is in fact based on a Discrete-Induced Transitions phenomenon (DIT).

DIT have already been studied in CRN models, more precisely in auto-catalytic CRNs:

$$S_i+S_j \rightharpoonup 2S_i, \qquad \emptyset \rightleftharpoons S_i, \qquad 1\leq i\neq j\leq m.$$

It is characterized by the fact that a limited number molecules of one chemical species can switch the entire bio-chemical state of a system, see Togashi and Kaneko [74], Bibbona et al. [13] and del Sole [22].

In our CRN with four chemical species and the initial state of the type

$$(0,N,0,\sqrt{N}),$$

it turns out that the growth of the fourth coordinate $(X_4(t))$ occurs only during time intervals whose duration are $O(1/\sqrt{N})$ and during them there is a large number of positive jumps of this process, of the order of $\sqrt{N}$. Recall that this phenomenon is only due to the law of mass action which drives the kinetics of the CRN. This is where boundary effects have a significant impact.

The switch effect occurs during these small time intervals, the occurrence of them is driven by the isolated instants of creation of particles of chemical species S_3. They play a critical role in the kinetics of the system, since this is at these instants, and only there, that the second coordinate $(X_2(t))$ can decrease.

We prove that $(X_4^N(t)/\sqrt{N})$ converges in distribution to a continuous time Markov jump process, the jumps being the consequence of the DIT. It is shown that the limit of the process is in fact an Additive Increase, Multiplicative Decrease (AIMD) process, with infinitesimal generator given by, for bounded function $f\in\mathcal{C}_1(\mathbb{R}_+)$ and $v\geq 0$,

$$\Omega_0(f)(v) = -\frac{1}{\gamma}vf'(v) + \int_0^{+\infty}\left(f\left(\sqrt{v^2+2\beta s}\right)-f(v)\right)e^{-s}\,\mathrm{d}s, \tag{1.14}$$

for some constants β, $\gamma>0$. The fact that a positive jump in the limit is a consequence of a large number of positive jumps of the initial process, in a small time interval, prevents us to use the classical J_1-Skorokhod topology. We prove the convergence in distribution for the M_1-Skorokhod topology instead, see Section 4.3.

The positive jumps of $X_4^N(t)$ are negative jumps of $X_2^N(t)$, and on the normal timescale, $O(\sqrt{N})$ happen. To see the decrease of $(X_2^N(t)/N)$, we accelerate the time, and look at the timescale $t\mapsto t\sqrt{N}$.

THEOREM. *If $X(0)=(0,N,0,v_N)$, with $(v_N/\sqrt{N})$ converging to $v\geq 0$, then the relation*

$$\lim_{N\to+\infty}\left(\frac{X_2\left(\sqrt{N}t\right)}{N}, t<t_\infty\right) = \left(\left(1-\frac{t}{t_\infty}\right)^2, t<t_\infty\right)$$

holds for the convergence in distribution, with

$$t_\infty = \sqrt{2}\,\frac{\kappa_5}{\kappa_0^2}\sqrt{\frac{\kappa_4}{\kappa_3}}\,\Gamma\left(\frac{\kappa_0}{2\kappa_5}\right)\Big/\,\Gamma\left(\frac{\kappa_5+\kappa_0}{2\kappa_0}\right),$$

where Γ *is the classical* Gamma *function.*

The quite technical proof of this theorem is done using multiple time changes. In particular, since the process $X_2^N(t)$ basically does not move if $X_3^N(t)=0$, the timescale $(\ell_1(t))$ defined by, for $t{\geq}0$,

$$L_1^N(t) \stackrel{\text{def.}}{=} \int_0^t \mathbb{1}_{\left\{U_3^N(s)\geq 1\right\}}\,\mathrm{d}s \quad \text{and} \quad \ell_1^N(t) \stackrel{\text{def.}}{=} \inf\{s{>}0 : L_1^N(s){>}t\}.$$

proves to be the relevant timescale to see the evolution of $(X_2^N(t)/N)$.

The process $(X_4^N(t\sqrt{N})/\sqrt{N})$ behaves like a fast process, and we prove that its occupation measure converges to the invariant measure of the AIMD, defined by Relation (1.14). This invariant measure is calculated explicitly in Section dedicated to the AIMD process.

5.3. Chapter 4: Analysis of stochastic chemical reaction networks with a hierarchy of timescales. This Chapter focus on a class of CRNs, that we call k-Unary CRNs.

k-*Unary Chemical Reaction Networks.* The only chemical reactions for this class of CRNs are as follows for $1{\leq}i{\neq}j{\leq}n$,

$$k_iS_i \xrightarrow{\kappa_{ij}} k_jS_j, \quad k_iS_i \xrightarrow{\kappa_{i0}} \emptyset, \quad \emptyset \xrightarrow{\kappa_{0i}} k_iS_i,$$

provided that, respectively, $\kappa_{ij}{>}0$, $\kappa_{i0}{>}0$, or $\kappa_{0i}{>}0$. The second reaction, resp. last reaction, is the spontaneous destruction, resp. creation, of k_i molecules of chemical species S_i. For $1{\leq}i{\leq}n$, it is assumed that the time evolution of the ith coordinate is a jump process whose jumps are $\pm k_i$. Here k_iS_i is the only *complex* involving the chemical species S_i.

The parameters of these networks is a matrix $R_\kappa{=}(\kappa_{ij}, 0{\leq}i,j{\leq}n)$ and a vector (k_i). They are quite general in fact, modulo an irreducibility property for the matrix R_κ. They can be described as transformations of sets of k_i molecules of S_i into k_j molecules of S_j, if $\kappa_{ij}{>}0$. In state $x{=}(x_j)$, the ith coordinate decreases at a rate proportional to $x_i^{(k_i)}$.

Scaling External Input Rates. The scaling investigated in this paper is as follows. For all $i{\in}\{1,\ldots,n\}$ such that $\kappa_{0i}{>}0$, the creation of chemical species S_i is scaled by N, it becomes

$$\emptyset \xrightarrow{N\kappa_{0i}} k_iS_i. \tag{1.15}$$

The other reaction rates do not change. Rather than starting from a "large" initial state, this scaling regime assume heavy traffic conditions at the entrance of the CRNs. A natural question in this setting is of establishing a limit theorem on the orders of magnitude in N of the coordinates of $(X_N(t)){=}(X_i^N(t))$. This scaling has already been considered in Togashi and Kaneko [73] for CRNs and probably in many other examples.

A basic example of such a situation is the k-Unary CRN with one chemical species,

$$\emptyset \underset{\mu}{\overset{\lambda N}{\rightleftharpoons}} k_1S_1.$$

It can be easily seen that, under convenient initial conditions, the scaled process

$$\left(\frac{X_1^N\left(t/N^{(1-1/k_1)}\right)}{N^{1/k_1}}\right) \tag{1.16}$$

converges in distribution to a non-trivial deterministic function, the solution of an ODE.

A Hierarchy of Timescales. Coming back to our CRNs, if we assume a kind of equilibrium of flows in the network, due to the external inputs of the order of N, the input flow through each node should be also of the same order of N. The case of the CRN with a single node suggests then that the state variable of the ith node $(X_i^N(t)/N^{1/k_i})$, $1{\leq}i{\leq}n$, should be of the order of N^{1/k_i}.

Relation (1.16) also states that the "natural" timescale of $(X_i(t)/N^{1/k_i})$ is $(t/N^{(1-1/k_i)})$. In particular, this implies that, at the "normal" timescale (t), all coordinates $(X_i(t)/N^{1/k_i})$ whose index $i{\in}\{1,\dots,n\}$ is such that $k_i{\geq}2$, are fast processes. The CRN exhibits in fact a hierarchy of timescales: The process associated to $(X_j(t)/N^{1/k_j})$ is faster than the process $(X_i(t)/N^{1/k_i})$ if $k_j{>}k_i$. A limit theorem to establish the convergence of the scaled process

$$\left(\frac{X_i^N(t)}{N^{1/k_i}}\right) \tag{1.17}$$

has to handle this multi-timescale feature and the interactions of the coordinates.

Outline of the Chapter. The goal of this Chapter is of establishing a limit theorem for the convergence in distribution of the scaled process (1.17) :

- For the occupation measure of the coordinates of the Markov process whose indices $i{\in}\{1,\dots,n\}$ are such that $k_i{\geq}2$;
- For the vector of the other components for the uniform topology.

The proof of this result is done in several steps. The identification of the limit of the sequence of occupation measures is the most challenging. The difficulty is to identify simultaneously all the species that evolve at the same timescale, using only one equation. Relative entropy functions associated to each timescale and convexity arguments are the main ingredients of the proof.

5.4. On the Recurrence Properties of Stochastic Chemical Reaction Networks with Two Species. This last chapter focus on a general class of weakly reversible stochastic chemical reaction networks (CRNs) with two chemical species S_1 and S_2. They are referred to as 2D-CRNs.

Our goal in this chapter is of giving an insight on the amount of time required for the process to return from a state in the interior of the domain to the boundary of it.

More precisely, we set $(X(t))$ the process associated to a weakly reversible 2D-CRN, and V the entropy function : for $x{=}(x_1,x_2){\in}\mathbb{R}_+^2$,

$$V(x) = v(x_1) + v(x_2) \quad \text{with} \quad v : y \in \mathbb{N} \mapsto y\ln(y){-}y{+}1,$$

with the convention $0\ln(0){=}0$.

Starting from a large state in the interior of $\mathbb{N}^2$, we show that there exists some $K{>}0$ such that

$$T_K \stackrel{\text{def.}}{=} \inf\{t{\geq}0 : \min(X_1(t),X_2(t)) \leq K\}$$

is integrable and that we can find a constant C_0 such that

$$\mathbb{E}_x(T_K) \leq C_0 V(x), \tag{1.18}$$

for all $x{=}(x_1,x_2)$ such that $\min(x_1,x_2){>}K$.

The proof is done with the use of Filonov's formulation of the Lyapunov condition introduced in Chapter 2 with the entropy function. One has to show that there exists positive constants γ, K and an integrable stopping time τ such that, for $x{=}(x_1,x_2){\in}\mathbb{N}^2$ with $\min(x_1,x_2){\geq}K$, then

$$\mathbb{E}_x(V(X(\tau)){-}V(X(0))) \leq -\gamma\mathbb{E}_x(\tau). \tag{1.19}$$

In this case, it can be shown that Relation (1.18) holds.

Filonov's approach has already been used for specific examples of CRNs in Chapter 2 and Chapter 3 to prove the positive recurrence of the associated Markov processes. In this chapter we show that it can also be used in the context of a quite large class of CRNs.

This result can also be useful to show the stability of such a CRN. The stability of this class of CRNs, i.e. the positive recurrence of $(X(t))$, is investigated in Agazzi et al. [2] by showing that the entropy function V is a Lyapunov function for the infinitesimal generator of $(X(t))$, i.e. it considers essentially the case when τ is taken as t_1. The proof of Filonov's criterion in the interior of the state space can be considered as an alternative to the difficult proof of Lyapunov's citerion in the same domain.

To establish Relation (1.19), a finite partition of the interior of the state space is used to define a convenient stopping time τ. Ideas related to the notion of endotactic CRN developed in [10] are used. For some subsets of the partition, taking $\tau{=}t_1$ is enough, where t_1 is the first instant of jump of the process $(X(t))$. For the other cases, functional limit theorems for the Markov process on a convenient timescale are necessary to define the appropriate τ.

1.A. Topologies in the Skorokhod space

I will here give a quick presentation of the four topologies used in my PhD thesis: the uniform topology, the J_1-Skorokhod topology, the M_1-Skorokhod topology and the S-Jakubowski topology. In this order, these topologies are presented from the strongest to the weakest, meaning that if a sequence converges for one of it, it converges for all the following ones. The topologies are presented on the set of càdlàg functions from $[0,1)$ to $\mathbb{R}$, noted $\mathcal{D}_{[0,1)}(\mathbb{R})$. The extension to $\mathcal{D}_{\mathbb{R}_+}(\mathbb{R}^m)$ can be found in Billingsley [14], Whitt [77].

See Section 11.6 of Whitt [77] for a sum up of this method.

All four topologies that we will present do not verify all the required properties. The uniform topology is not separable, but its criterion for compactness is very useful. Both J_1 and M_1-Skorokhod topologies satisfy all these properties, but the S-Jakubowski topology cannot be metricized. However, it is constructed so that the Prokhorov theorem still holds. We will give a characterization of the compact sets for each topologies.

1.A.1. The uniform topology. The topology associated to the uniform norm $\|-\|_\infty$ is a priori used for the set of continuous functions $\mathcal{C}([0,1],\mathbb{R}_+)$.

It can be used for a sequence of processes of $\mathcal{D}_{[0,1)}(\mathbb{R})$ when its limit is a *continuous* process. It is the most commonly used topology when the sequence converges to the solution of an ODE, as in the classical scaling of Section 4.2.3. For this topology, the criterion for compactness is the same as the one from Arzelà-Ascoli's theorem:
A set $K \subset \mathcal{D}_{[0,1)}(\mathbb{R})$ of càdlàg functions is relatively compact for the uniform topology if it is uniformly bounded at 0, and uniformly equicontinuous, i.e. if

$$\sup_{x\in K} |x(0)| < +\infty,$$

and defining the *modulus of continuity* ω as

$$\omega_\delta(x) \overset{\text{def.}}{=} \sup_{|t-s|<\delta} |x(t)-x(s)|$$

for $\delta > 0$ and $x \in \mathcal{D}_{[0,1)}(\mathbb{R})$, if

$$\lim_{\delta\to 0} \sup_{x\in K} \omega_\delta(x) = 0.$$

It can be shown that a sequence of processes tight for the uniform topology is indeed converging in distribution, still for the uniform topology. However, the topological space associated to this topology *is not separable*: the set of functions $\{\mathbb{1}_{\{[x,1)\}}, x \in \mathbb{R}\}$ constitute an uncountable set of functions all at distance 1 of one another, which prevents the use the Prokhorov theorem.

When the limit of the sequence of processes presents a discontinuity, this topology is not weak enough to be used. Indeed, it lacks a tolerance on the small perturbation of the time. For example, the functions $\mathbb{1}_{\{[1/2-1/N,1)\}}$ with graph in Figure 3 do not converge to $\mathbb{1}_{\{[1/2,1)\}}$ for the uniform topology.

The uniform topology is stronger than all the other topologies presented below, and therefore a sequence converging for the uniform topology will converge for the three other topologies presented here.

1.A.2. The J_1-Skorokhod topology. The J_1-Skorokhod topology is the most commonly used when working on the Skorokhod space $\mathcal{D}_{[0,1)}(\mathbb{R})$. In this topology, the sequence $\mathbb{1}_{\{[1/2-1/N,1)\}}$ of Figure 3 indeed converges to the function $\mathbb{1}_{\{[1/2,1)\}}$. The idea is to add the missing flexibility on the time variable, in particular flexibility in the location of jumps, using a time change. A distance can be defined to have a complete, metric space, see Whitt [77].

On this space can be defined a criterion for compactness similar to the Arzelà-Ascoli theorem. We introduce the *modified* modulus of continuity,

$$\omega'_\delta(x) = \inf\left\{\max_{i\leq r}\rho(x,[t_{i-1},t_i)) : 0 = t_0 < \ldots < t_r = T, \inf(t_i - t_{i-1}) \geq \delta\right\}$$
$$\text{with} \quad \rho(x,I) = \sup_{s,t\in I}|x(s) - x(t)|$$

for $\theta > 0$, $x \in \mathcal{D}_{[0,1)}(\mathbb{R})$, and $I \subset [0,T]$ an interval. A set $K \subset \mathcal{D}_{[0,1)}(\mathbb{R})$ of càdlàg functions is relatively compact in J_1 if it is uniformly bounded on $[0,1)$, and "piecewise uniformly equicontinuous", i.e. if

$$\sup_{x\in K}\|x\|_\infty < +\infty,$$

and if

$$\lim_{\delta\to 0}\sup_{x\in K}\omega'_\delta(x) = 0.$$

Intuitively, this modulus of continuity allows a finite number of jumps at the instants t_i, but between these, the process must be uniformly equicontinuous. In particular, this topology requires that the converging functions have jumps corresponding to each jump in the limit function. In particular, the sequence of functions defined by Figure 4 does not converges in the J_1-Skorokhod topology. In the framework of CRNs, the process studied is often scaled in space as $(X_N(t)/N^\alpha)$, for some $\alpha > 0$. Therefore, if we want to show the converges in distribution of this scaled process to a process with jumps, the jump should be a consequence of a large number of small jumps, happening in a time interval vanishing when N gets large. Therefore, this topology is not "weak enough" for our purpose.

1.A.3. The M_1-Skorokhod topology. To remedy to this issue, a strategy is to extend the ε-balls around a sample path. Indeed, the ε-ball around the function $\mathbb{1}_{\{[1/2,1)\}}$ in the J_1-topology only extends a little the uniform ε-ball to the left and to the right, along the "time dimension", see Figure 7. Adding all the points connecting the left and right limits of the jumps, we get an ε-ball as in Figure 8, which allows more flexibility to the convergent sequences. This is formally carried out by introducing the *complete graph* of $x \in \mathcal{D}_{[0,1)}(\mathbb{R})$:

$$\Lambda(x) \stackrel{\text{def.}}{=} \{(t,z) \in [0,1)\times\mathbb{R} : \ z = \alpha x(t-) + (1-\alpha)x(t), \quad \alpha \in [0,1]\},$$

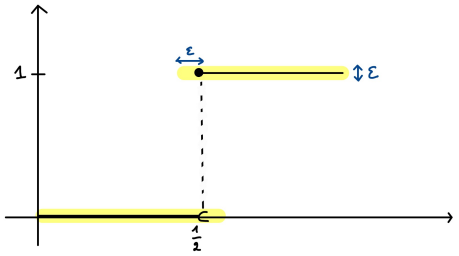


FIGURE 7. ε-ball in the J_1 topology

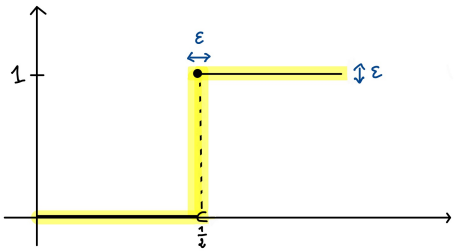


FIGURE 8. ε-ball in the M_1 topology

and setting the distance in M_1 as some kind of distance between complete graphs (involving a parametrization and an ordering of the complete graphs). See Whitt [77] for a formal definition. The topology associated to this distance is complete, separable, weaker than the J_1-Skorokhod topology, and this time the sequence of Figure 4 converges to $\mathbb{1}_{\{[1/2,1)\}}$ in this topology.

Besides, one can define a criterion for compactness in this topology, using another modulus of continuity : for $x \in \mathcal{D}_{[0,1)}(\mathbb{R})$, $0 < \delta < 1$,

$$\omega''_\delta(x) \overset{\text{def.}}{=} \sup_{t\in[0,1-\delta]} \sup_{t\leq t_1<t_2<t_3<t+\delta} d(x(t_2),[x(t_1),x(t_3)]),$$

where $d(\cdot, A)$ stands for the distance to the set A.

A set $K \subset \mathcal{D}_{[0,1)}(\mathbb{R})$ of càdlàg functions is relatively compact in M_1 if it is uniformly bounded on $[0,1)$, and i.e. if

$$\sup_{x\in K} \|x\|_\infty < +\infty,$$

and if

$$\lim_{\delta\to 0} \sup_{x\in K} \omega''_\delta(x) = 0.$$

1.A.4. The S-Jakubowski topology. In the example of Figure 5, the function presents a "spike" at $t \approx 1/4$, whose length goes to zero when n gets large. We could hope for the limit of this sequence to be $\mathbb{1}_{\{[1/2,1)\}}$, but this sequence does not converges in the M_1-Skorokhod topology. However, the S-Jakubowski topology can be used to study this kind of sequence.

This last topology is the least intuitive one. Indeed, contrary to the others, it is not constructed on a the definition of a distance, and even *cannot be metricized*. This topology has been introduced to match with the following criterion for compact set : if $K \subset \mathcal{D}_{[0,1]}(\mathbb{R})$ is uniformly bounded, i.e. verify

$$\sup_{x\in K} \|x\|_\infty < C < +\infty,$$

and if for all $a < b$,

$$\sup_{x\in K} N^{a,b}(x) < +\infty$$

where $N^{a,b}(x)$ is the number of up-crossing given levels $a < b$ of a function x, (i.e. as the maximal $k \geq 1$ such that there are $0 \leq t_1 < t_2 \ldots < t_{2k} \leq 1$ that verify $x(t_{2i-1}) < a$ and $x(t_{2i}) > b$ for all $1 \leq i \leq k$), then we want the set K to be relatively compact for the topology chosen. The construction, and more details on this topology can be found in Jakubowski [44], we will call this topology the S-Jakubowski topology. It is a *sequential* topology, which is separable, and *weaker* than the J_1 and M_1-Skorokhod topologies. Even if the topological space

cannot be metricized, the topology is constructed in such a way that an equivalent of Prokhorov's theorem holds, and knowing the tightness of a sequence of processes still implies its relative compactness. To sum up, to show the convergence in distribution of a sequence (Y_n) of processes for the S-Jakubowski topology, one has to show

- its tightness, by showing that $\{\|Y_n\|_\infty\}$ and $\{N^{a,b}(Y_n), n \geq 1\}$ are uniformly tight, for all $a < b$.
- that the finite-dimensional distributions (on a dense subset of $[0,1]$) converge in distribution.

One of the advantages of this topology is that, contrary to the J_1 and M_1-Skorokhod topologies, the sum of two converging sequences still converges to the sum of their limits, without making any assumptions on the time instants of jumps of the limits.

1.B. Proof of the Deficiency zero theorems

1.B.1. Deterministic theorem. The theorem proven here is the following :

THEOREM 1.4. *Let $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$ be a weakly reversible CRN with deficiency zero. For any constant set of positive constant of reactions $(\kappa_r, r \in \mathcal{R})$, the solution of the ODE associated to the deterministic CRN with mass action kinetic $\{\mathcal{S}, \mathcal{C}, \mathcal{R}, \kappa\}$ has precisely one equilibrium in each of its positive stoichiometric compatibility class. Besides, that equilibrium is asymptotically locally stable.*

Note here that this is a "weak" version of the Deficiency zero Theorem.

The proof given here essentially follows Lectures 4 and 5 of Feinberg [25] and Gunawardena [36].

1.B.2. General notations. In the following proof, the set of positive real numbers is noted $\mathbb{P}$. The number of element of a finite set I is noted $|I|$.

The concentrations of species are elements of in $\mathbb{P}^{\mathcal{S}}$ and the complexes, elements of $\mathbb{N}^{\mathcal{C}}$. The *support* of a vector $x \in \mathbb{R}^{\mathcal{C}}$ (or $\mathbb{R}^{\mathcal{S}}$) is the set

$$\mathrm{supp}(x) \stackrel{\text{def.}}{=} \{y \in \mathcal{C} : x_y \neq 0\}.$$

We will need a few standard notations from linear algebra. For a finite set of indices I (here, $I = \mathcal{C}$ or $\mathcal{S}$), $\mathbb{R}^I$ is a vector space, on which we define the usual inner product $\langle \cdot, \cdot \rangle$: for $x, y \in \mathbb{R}^I$,

$$\langle x, y \rangle = \sum_{i \in I} x_i y_i.$$

The canonical basis of $\mathbb{R}^{\mathcal{C}}$ is noted $(\omega_y,\ y \in \mathcal{C})$.

For a set $A \subset \mathbb{R}^I$, we note Span A the linear vector space generated by A.

The *dimension* of H a linear subspace of $\mathbb{R}^I$ is noted $\dim H$. $H^\perp$ is the set of vectors that are orthogonal to all elements of H.

If E and F are two euclidean spaces, and g is a linear map, $g : E \to F$, we denote by Ker g the *kernel* of g, Im g the *image* of g. The rank theorem states that

$$\dim(\mathrm{Ker}\ g) + \dim(\mathrm{Im}\ g) = \dim(E).$$

Finally, the transpose of the linear map g is noted $g^T : F \to E$.

1.B.3. Model. We recall the ODE (1.2) associated to the CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \kappa)$:

$$\dot{c} = f(c) = \sum_{r = y^- \rightharpoonup y^+ \in \mathcal{R}} \kappa_r c^{y^-} (y^+ - y^-). \tag{1.20}$$

The function f can be decomposed in $f = Y A_\kappa \psi$ with Y the canonical projection from $\mathbb{R}^{\mathcal{C}}$ to $\mathbb{R}^{\mathcal{S}}$: for all $y \in \mathcal{C}$, $Y(\omega_y) = y$, $A_\kappa : \mathbb{P}^{\mathcal{C}} \to \mathbb{P}^{\mathcal{C}}$ defined as

$$A_\kappa : x \mapsto \sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} \kappa_r x_{y^-} (\omega_{y^+} - \omega_{y^-}) \tag{1.21}$$

and $\psi : \mathbb{P}^{\mathcal{S}} \to \mathbb{P}^{\mathcal{C}}$ defined as

$$\psi : c \mapsto \sum_{y \in \mathcal{C}} c^y \omega_y. \tag{1.22}$$

Note that Y and A_κ are linear.

Finally, we define

$$S = \text{Span}\,\{y^+ - y^- : y^- \to y^+ \in \mathcal{R}\}. \tag{1.23}$$

The positive stoichiometric compatibility class of a $c \in \mathbb{P}^{\mathcal{S}}$ is the set $\mathbb{P}^{\mathcal{S}} \cap (S + c)$.

1.B.4. The kernel of A_κ. The following proposition specify the form of the kernel of A_κ.

PROPOSITION 1.5. *Let $\{\mathcal{S}, \mathcal{C}, \mathcal{R}, \kappa\}$ be a* weakly reversible *deterministic CRN with mass action kinetic. Let $\{\Lambda_1, \ldots, \Lambda_\ell\}$ be its linkage classes. Then the kernel of A_κ has a basis $\{\chi_1, \ldots, \chi_\ell\}$, and for all $1 \leq i \leq \ell$,*

$$\text{supp}\chi_i = \Lambda_i \quad \textit{and} \quad \chi_i \in (\mathbb{R}_+)^{\mathcal{C}}. \tag{1.24}$$

The original result, that can be found in Feinberg [25], states that the vectors of the basis of Ker A_κ are supported by the strong *terminal* linkage classes (see Feinberg [25] for the definition of terminal linkage classes), but with a CRN weakly reversible, the linkage classes and terminal linkage classes are identical.

The following proof of the proposition is a condensed version of the proof found in the Appendix of [29].

PROOF. First, lets show that if $\chi \in \text{Ker}\ A_\kappa$, then $|\chi| \in \text{Ker}\ A_\kappa$. Let $\chi \in \text{Ker}\ A_\kappa$. For $y_0 \in \mathcal{C}$, looking at the coordinates in y_0 of $A_\kappa(\chi)$, we have:

$$\sum_{r=y_0 \rightharpoonup y^+ \in \mathcal{R}} \kappa_r \chi_{y_0} = \sum_{r=y^- \rightharpoonup y_0 \in \mathcal{R}} \kappa_r \chi_{y^-}. \tag{1.25}$$

Hence, for all $y_0 \in \mathcal{C}$,

$$\sum_{r=y_0 \rightharpoonup y^+ \in \mathcal{R}} \kappa_r \left|\chi_{y_0}\right| \leq \sum_{r=y^- \rightharpoonup y_0 \in \mathcal{R}} \kappa_r \left|\chi_{y^-}\right|, \tag{1.26}$$

and then summing on the $y_0 \in \mathcal{C}$, we get:

$$\sum_{y_0 \in \mathcal{C}} \sum_{r=y_0 \rightharpoonup y^+ \in \mathcal{R}} \kappa_r \left|\chi_{y_0}\right| \leq \sum_{y_0 \in \mathcal{C}} \sum_{r=y^- \rightharpoonup y_0 \in \mathcal{R}} \kappa_r \left|\chi_{y^-}\right|,$$

but this inequality is in fact an equality (it is just a different partition on the set of reactions), and for all $y_0 \in \mathcal{C}$, inequality (1.26) is an equality, and $A_\kappa(|\chi|) = 0$.

We can now make the assumption that every $\chi \in \text{Ker}\ A_\kappa$ we use has non negative coefficients.

Using Relation (1.25), we can show that for $\chi \in \text{Ker}\ A_\kappa$, χ non negative, if $y, y' \in \mathcal{C}$ are in the same linkage class and $\chi(y) = 0$, then $\chi(y') = 0$ (and therefore χ is zero on all the linkage class of y). This specific result is closely linked to the weak reversibility of the CRN, and therefore of each linkage class.

Setting $L_i = \text{Span}\ \Lambda_i$, we have $\mathbb{R}^{\mathcal{C}} = L_1 \oplus \ldots \oplus L_l$ and for all $1 \leq i \leq \ell$, L_i is invariant under A_κ (i.e. if $x \in L_i$, $A_\kappa(x) \in L_i$). We can therefore find a basis of Ker A_κ adapted to the direct sum.

Let $1 \leq i \leq \ell$. Seeing that $\sum_{y \in \Lambda_i} \omega_y \in (\text{Im}\ A_\kappa)^\perp \cap L_i$, we know that

$$\dim((\text{Im}\ A_\kappa) \cap L_i) < \dim(L_i),$$

and since L_i is invariant under A_κ, we have

$$\dim((\text{Ker } A_\kappa) \cap L_i) = \dim L_i - \dim((\text{Im } A_\kappa) \cap L_i) > 0,$$

and therefore Ker $A_\kappa \cap L_i \neq \{0\}$. The basis of Ker A_κ contains at least one vector for each linkage class.

Remains to show that it contains *only* one by linkage class, i.e. that for $1 \leq i \leq \ell$, if $\chi, \chi' \in$ Ker $A_k \cap L_i$, then $\chi = \alpha\chi'$ for some $\alpha \in \mathbb{R}$, which is straightforward using the the previous statements. □

1.B.5. Another definition of the deficiency. At this state of the proof, we can show that the deficiency can be understood as the dimension of a linear space. This proves that the deficiency is always a non negative integer. Note that the following proposition only holds when the CRN is weakly reversible, when both terminal linkage classes and linkage classes are identical.

PROPOSITION 1.6. *Recall that the* deficiency *of a CRN is defined (see Section 2.2) as*

$$\delta = |\mathcal{C}| - \dim(S) - \ell$$

where ℓ is the number of linkage classes of the CRN.

If the CRN is weakly reversible, the following Relation holds:

$$\delta = \dim(\text{Ker } Y \cap \text{Im } A_\kappa). \tag{1.27}$$

PROOF. For this proof, we need to introduce the equivalent of the set S in the space of the complexes:

$$T \overset{\text{def.}}{=} \text{Span } \left\{\omega_{y^+} - \omega_{y^-},\ y^- \rightharpoonup y^+ \in \mathcal{R}\right\} \tag{1.28}$$

where (ω_y) is the canonical basis of $\mathbb{R}^{\mathcal{C}}$.

Straightforwardly, Im $A_\kappa \subset T$. We want to show that when the CRN is weakly reversible, Im $A_\kappa = T$. We will proceed by equality of dimension. Proposition 1.5 and the rank theorem gives $\dim(\text{Im } A_\kappa) = |\mathcal{C}| - \ell$. Lets look at the dimension of T. First, setting for all $1 \leq i \leq \ell$, $T_i \overset{\text{def.}}{=} \text{Span } \left\{\omega_{y^+} - \omega_{y^-}, y^- \rightharpoonup y^+ \in \mathcal{R}, y^-, y^+ \in \Lambda_i\right\}$, it is straightforward to show that

$$T = T_1 \oplus \ldots \oplus T_\ell.$$

It is therefore sufficient to find the dimension of each T_i. Let $1 \leq i \leq \ell$. If $|\Lambda_i| = p_i$ and $\Lambda_i = \{y_1, \ldots, y_{p_i}\}$, then the set

$$(\omega_{y_2} - \omega_{y_1}, \omega_{y_3} - \omega_{y_1}, \ldots, \omega_{y_{p_i}} - \omega_{y_1})$$

is a basis of T_i, and therefore $\dim(T_i) = p_i - 1$. Hence,

$$\dim(T) = \sum_{i=1}^{\ell} \dim(T_i) = \sum_{i=1}^{\ell} (p_i - 1) = |\mathcal{C}| - \ell,$$

and Im $A_\kappa = T$. To conclude, first we notice that the restriction of Y to T, noted $Y_{|T}$, is surjective on S, i.e. that $\dim(\text{Im } Y_{|T}) = \dim(S)$. We then have :

$$(\text{Ker } Y \cap \text{Im } A_\kappa) = (\text{Ker } Y \cap T) = \text{Ker } Y_{|T}$$

and therefore,

$$\dim(\text{Ker } Y \cap \text{Im } A_\kappa) = \dim(\text{Ker } Y_{|T}) = \dim(T) - s = n - \ell - s = \delta.$$

The second equation is given by the rank theorem. □

1.B.6. Existence of fixed points.

PROPOSITION 1.7. *If a chemical reaction network has deficiency zero and is weakly reversible, then for some $c \in \mathbb{P}^{\mathcal{S}}$, the relation $A_k\psi(c) = 0$ holds.*

This is not the strongest result. It can be shown that if the CRN has deficiency zero, a positive equilibrium exists *if and only if* it is weakly reversible.

PROOF. We set $\omega_{\Lambda_i} = \sum_{y\in\Lambda_i} \omega_y$, and the set

$$U \overset{\text{def.}}{=} \text{Im } Y^T + \text{Span } (\omega_{\Lambda_1}, \dots, \omega_{\Lambda_l}) .$$

Lets show that $U = \mathbb{R}^{\mathcal{C}}$.

The ω_{Λ_i} are linearly independent, therefore, $\dim(\text{Span } (\omega_{\Lambda_1}, \dots, \omega_{\Lambda_\ell})) = \ell$. Using the explicit expression of A_κ, it is straightforward to show that for $1 \leq i \leq \ell$, $\omega_{\Lambda_i} \in (\text{Im } A_\kappa)^\perp$. Therefore, Span $(\omega_{\Lambda_1}, \dots, \omega_{\Lambda_l}) \subset (\text{Im } A_\kappa)^\perp$, and using Proposition 1.5, by equality of their dimensions, we have

$$(\text{Im } A_\kappa)^\perp = \text{Span } (\omega_{\Lambda_1}, \dots, \omega_{\Lambda_l}) .$$

Using two elementary results of linear algebra and the last Relation, we get

$$U = (\text{Ker } Y)^\perp + (\text{Im } A_\kappa)^\perp = (\text{Ker } Y \cap \text{Im } A_\kappa)^\perp,$$

and since the deficiency $\delta = \dim(\text{Ker } Y \cap \text{Im } A_\kappa)$ is zero, we indeed have $U = \mathbb{R}^{\mathcal{C}}$.

Now lets find a $c \in \mathbb{P}^{\mathcal{S}}$ such that $A_\kappa\psi(c) = 0$. The function log on $\mathbb{P}^{\mathcal{S}}$ (or $\mathbb{P}^{\mathcal{C}}$) is to be understood component wise: for $c \in \mathbb{P}^{\mathcal{S}}$, $\log(c) = (\log(c_i), i \in \mathcal{S})$. For $c \in \mathbb{P}^{\mathcal{S}}$, $\psi(c) \in \mathbb{P}^{\mathcal{C}}$ and

$$\log(\psi(c)) = Y^T \log(c). \tag{1.29}$$

Since $U = \mathbb{R}^{\mathcal{C}}$, we can find some $x \in \mathbb{R}^{\mathcal{C}}$ and $\zeta \in \mathbb{R}^\ell$ such that

$$\log\left(\sum_{i=1}^{l} \chi_i\right) = Y^T x - \sum_{i=1}^{l} \zeta_i \omega_{\Lambda_i}$$

where the χ_i are the vectors of the basis of Ker A_κ, introduced in Proposition 1.5. Setting $c = e^x > 0$ (component wise) and $\lambda_i = e^{\zeta_i}$, and using (1.29) we get:

$$\begin{aligned}\log\psi(c) = Y^T \log(c) &= \sum_{i=1}^{l} \sum_{y\in\Lambda_i} \left[\log(\langle\chi_i, \omega_y\rangle) + \log(\lambda_i)\right] \omega_y \\ &= \log\left(\sum_{i=1}^{l} \lambda_i\chi_i\right)\end{aligned}$$

(χ_i) being a basis of Ker A_κ, $\psi(c) \in$ Ker A_κ and $c > 0$ by construction, so we have found an equilibrium. All the calculus with the log function work because each complex $y \in \mathcal{C}$ is present in one and only one of the χ_i, with a positive coefficient. □

We will see that knowing the existence of one fixed point helps to show the existence of the ones in other positive stoichiometric compatibility classes.

1.B.7. Fixed points for which $A_k\psi(c) = 0$.

PROPOSITION 1.8. *We set*

$$Z \overset{\text{def.}}{=} \{c \in \mathbb{P}^{\mathcal{C}} : A_k\psi(c) = 0\}.$$

Each stoichiometric compatibility class meets Z on exactly one state.

First, lets show the following lemma :

LEMMA 1.9. *The following equality holds :*

$$Z = \{c \in \mathbb{P}^{\mathcal{S}} : \log c - \log c^* \in S^\perp\}$$

where S is defined by Relation (1.23),

PROOF OF THE LEMMA. Let $c \in Z$, $y^- \rightharpoonup y^+ \in \mathcal{R}$, and $j \in \{1, \dots, \ell\}$ such that $y^-, y^+ \in \Lambda_j$. We have to show that $\langle \log(c) - \log(c^*), y^+ - y^- \rangle = 0$.

Since $\psi(c), \psi(c^*) \in \operatorname{Ker} A_\kappa$, we can write

$$\psi(c) = \sum_{i=0}^{l} \lambda_i \chi_i \quad \text{and} \quad \psi(c^*) = \sum_{i=0}^{l} \mu_i \chi_i,$$

for some $(\lambda_i), (\mu_i) \in (\mathbb{R}_+^*)^\ell$.

Therefore,

$$\begin{aligned}
\left\langle \log(c) - \log(c^*), y^+ - y^- \right\rangle &= \left\langle \log(c) - \log(c^*), Y(\omega_{y^+} - \omega_{y^-}) \right\rangle \\
&= \left\langle Y^T \log(c) - Y^T \log(c^*), \omega_{y^+} - \omega_{y^-} \right\rangle \\
&= \left\langle \log \psi(c) - \log \psi(c^*), \omega_{y^+} - \omega_{y^-} \right\rangle \\
&= \left\langle \log \psi(c), \omega_{y^+} - \omega_{y^-} \right\rangle - \left\langle \log \psi(c^*), \omega_{y^+} - \omega_{y^-} \right\rangle \\
&= \log(\lambda_j \left\langle \chi_j, \omega_{y^-} \right\rangle) - \log(\lambda_j \left\langle \chi_j, \omega_{y^+} \right\rangle) - \log(\mu_j \left\langle \chi_j, \omega_{y^-} \right\rangle) + \log(\mu_j \left\langle \chi_j, \omega_{y^+} \right\rangle) \\
&= 0
\end{aligned}$$

and therefore $Z \subset \{c \in \mathbb{P}^{\mathcal{S}} : \log c - \log c^* \in S^\perp\}$.

Let $c = c^* e^u$ with $u \in S^\perp$. We want to show that $A_\kappa \psi(c) = 0$. Let $y_0 \in \mathcal{C}$. If $y^- \rightharpoonup y_0 \in \mathcal{R}$, then $\langle u, y^- \rangle = \langle u, y_0 \rangle$ since $u \in S^\perp$.

$$\begin{aligned}
\langle A_\kappa \psi(c), \omega_{y_0} \rangle &= \sum_{r = y^- \rightharpoonup y_0 \in \mathcal{R}} k_r c^{y^-} - \sum_{r = y_0 \rightharpoonup y^+ \in \mathcal{R}} \kappa_r c^{y_0} \\
&= \sum_{r = y^- \rightharpoonup y_0 \in \mathcal{R}} \kappa_r (c^*)^{y^-} e^{\langle u, y^- \rangle} - \sum_{r = y_0 \rightharpoonup y^+ \in \mathcal{R}} \kappa_r (c^*)^{y_0} e^{\langle u, y_0 \rangle} \\
&= e^{\langle u, y_0 \rangle} \langle A_\kappa \psi(c^*), \omega_{y_0} \rangle \\
&= 0
\end{aligned}$$

Which gives the other inclusion. □

PROOF OF PROPOSITION 1.8. Lets take two equilibria $c, c' \in Z$ in the same compatibility class: $c' - c \in S$. We set $u, v \in S^\perp$ such that $c = c^* e^u$ and $c' = c^* e^v$. We have

$$\langle c^*(e^u - e^v), u - v \rangle = 0 = \sum_{y \in \mathcal{S}} c_y^* (e^{u_y} - e^{v_y})(u_y - v_y),$$

and each real of the sum is non negative. Therefore, all of them must be null. Since $c^* \in \mathbb{P}^{\mathcal{S}}$, we must have $u = v$, which proves the uniqueness.

We now have to show that Z actually meets every stoichiometric compatibility class. Let $b \in \mathbb{P}^{\mathcal{S}}$. We want to find $u \in S^\perp$ such that $c^* e^u - b \in S$. To do so, we define the function $\phi : x \in \mathbb{R}^{\mathcal{S}} \mapsto \sum_{s \in \mathcal{S}} c_s^* e^{x_s} - b_s x_s$. Its gradient is given by:

$$\nabla \phi(x) = c^* e^x - b$$

The function ϕ is continuous, and goes to infinity when $\|x\| \to \infty$. It is strictly convex (since the function exponential is strictly convex and $y \mapsto -y$ is convex on $\mathbb{R}$).

Now let $\bar{\phi} : S^\perp \to \mathbb{R}$ be the restriction of ϕ to $S^\perp$. $\bar{\phi}$ is still continuous and convex therefore the following set is compact:

$$C = \{x \in S^\perp : \bar{\phi}(x) \leq \bar{\phi}(0)\}.$$

We can find $u \in S^\perp$ such that for all $x \in S^\perp$, $\bar{\phi}(u) \leq \bar{\phi}(x)$. Thus, for all $\gamma \in S^\perp$, and at $\theta = 0$:

$$\begin{aligned} 0 &= \frac{d}{d\theta}\bar{\phi}(u + \theta\gamma) \\ &= \frac{d}{d\theta}\phi(u + \theta\gamma) \\ &= \nabla\phi(u) \cdot \gamma \end{aligned}$$

It follows that $\nabla\phi(u)$ is orthogonal to $S^\perp$ and therefore is in S, and $c^* e^u$ is the unique equilibrium in the stoichiometric compatibility class of b. □

1.B.8. Stability of the equilibrium. Let $c \in Z$. We are going to show that c is stable, using a Lyapunov function. Here, we say that c is stable if it is stable in its stoichiometric compatibility class only (since a trajectory can only stay in one stoichiometric class). We will look at the Lyapunov function on $c + S$. Let $V : \mathbb{P}^{\mathcal{S}} \to \mathbb{R}$ defined as:

$$V(x) = \sum_{s \in \mathcal{S}} x_s(\log x_s - \log c_s - 1) + c_s$$

With an elementary study of the function V, we can see that V is non negative, and that $h(x) = 0$ if and only if $x = c$. Besides, $\nabla h(x) = \log(x) - \log(c)$.

We have to show that if $(x(t))$ is the trajectory coming from $x_0 \in S + c$, then for all $t \geq 0$,

$$\frac{d}{dt}h(x(t)) = \langle \nabla h(x(t)), f(x(t)) \rangle \leq 0$$

where f is defined in Relation (1.20). We are going to show that for $x \in c + S$, we have $\nabla h(x) \cdot f(x) \leq 0$ with equality only at $x = c$. We will use the following inequality: for $\alpha, \alpha' \geq 0$,

$$e^\alpha(\alpha' - \alpha) \leq e^{\alpha'} - e^\alpha$$

(with equality when $\alpha = \alpha'$)

$$\begin{aligned} \langle \nabla h(x), f(x) \rangle &= \sum_{r = y^- \to y^+} \kappa_r x^{y^-} \left\langle (y^+ - y^-), (\log x - \log c) \right\rangle \\ &= \sum_{r = y^- \to y^+} \kappa_r c^{y^-} e^{\langle y^-, \log x - \log c \rangle} \left\langle (y^+ - y^-), (\log x - \log c) \right\rangle \\ &\leq \sum_{r = y^- \to y^+} \kappa_r c^{y^-} (e^{\langle y^-, \log x - \log c \rangle} - e^{\langle y^+, \log x - \log c \rangle}) \\ &= \left\langle \sum_{r = y^- \to y^+} \kappa_r c^{y^-} (\omega_{y^+} - \omega_{y^-}), \sum_{y' \in \mathcal{C}} e^{\langle y', \log x - \log c \rangle} \omega_{y'} \right\rangle \\ &= 0 \end{aligned}$$

We recognize $A_\kappa \psi(c)$ in the first sum of the third equality, which is 0. We have equality in $\langle \nabla h(x), f(x) \rangle \leq 0$ if for all $y^- \rightharpoonup y^+ \in \mathcal{R}$, $\langle \log x - \log c, y^+ - y^- \rangle = 0$, i.e. if $\log x - \log c \in S^\perp$, and since $Z \cap (S + c) = \{c\}$, this is equivalent to $x = c$. This proves the function h is a Lyapunov function, and c is a stable equilibrium.

1.B.9. Stochastic theorem. The proof of Theorem 1.2 is based on Anderson et al. [7]. It consist only on verifying that the probability measure defined by Relation (1.9) is an invariant distribution. We check that the probability measure defined by Relation (1.9) verify for all $x \in \mathcal{E}_0$,

$$\sum_{r = y^- \rightharpoonup y^+} \pi_{\mathcal{E}_0}(x - y^+ + y^-)\lambda_r(x + y^+ - y^-) = \sum_{r = y^- \rightharpoonup y^+} \pi_{\mathcal{E}_0}(x)\lambda_r(x).$$

We start from the left side of the equation, and we part the sum according to the value of the complex y^+. We then apply the complex balanced equation verified by c. Since c verify $A_\kappa(\psi(c)) = 0$, for any complex y_0,

$$\sum_{r=y^-\to y_0\in\mathcal{R}} \kappa_r c^{y^-} = c^{y_0} \sum_{r=y_0\to y^+} \kappa_r.$$

For $x, x' \in \mathbb{N}^m$ and $y \in \mathcal{C}$, $x \geq x'$ means that for all i, $x_i \geq x'_i$, and we will use the notation

$$x! = \prod_{i=1}^m x_i! \quad \text{and} \quad x^{(y)} = \frac{x!}{(x-y)!}\mathbb{1}_{\{x\geq y\}}.$$

For $x \in \mathcal{E}_0$,

$$\begin{aligned}
&\sum_{r=y^-\rightharpoonup y^+} \pi_{\mathcal{E}_0}(x - y^+ + y^-)\lambda_r(x + y^+ - y^-) \\
&\qquad = \sum_{y_0\in\mathcal{C}} \frac{c^{x-y_0}x^{(y_0)}}{x!} \sum_{r=y^-\rightharpoonup y_0\in\mathcal{R}} \kappa_r c^{y^-}\mathbb{1}_{\{x-y_0+y^-\geq y^-\}} \\
&\qquad = \sum_{y_0\in\mathcal{C}} \frac{c^{x-y_0}x^{(y_0)}}{x!} c^{y_0}\mathbb{1}_{\{x\geq y_0\}} \sum_{r=y_0\rightharpoonup y^+\in\mathcal{R}} \kappa_r \\
&\qquad = \sum_{r=y^-\rightharpoonup y^+} \pi_{\mathcal{E}_0}(x)\lambda_r(x).
\end{aligned}$$

CHAPTER 2

A Scaling Approach to Stochastic Chemical Reaction Networks

Contents

1. Introduction

This paper investigates the asymptotic properties of Markov processes associated to chemical reaction networks (CRNs). The state space of these processes is a subset of $\mathbb{N}^n$, where $n \geq 1$ is the number of chemical species. A chemical reaction is associated to a transition of the associated Markov process $(X(t))$ of the form, for $x=(x_i)\in\mathbb{N}^n$

$$x \longrightarrow x + \sum_1^n \left(y_i^+ - y_i^-\right) e_i,$$

where, for $i\in\{1,\dots,n\}$, e_i is the ith unit vector of $\mathbb{N}^n$ and y^+, $y^-\in\mathbb{N}^n$ that form the reaction $y^- \rightharpoonup y^+$. The kinetics used classically for CRNs is the *law of mass action*. This is expressed by the fact that the above transition has a rate proportional to

$$\frac{x!}{(x-y^-)!} \overset{\text{def.}}{=} \prod_{i=1}^n \frac{x_i!}{(x_i-y_i^-)!} = \prod_{i=1}^n x_i(x_i-1)\cdots(x_i-y_i^-+1),$$

provided that $x_i \geq y_i^-$ holds for all $i\in\{1,\dots,n\}$, it is 0 otherwise. A chemical reactions requires a minimal number of copies of some chemical species to take place.

An important feature of these processes is that this transition rate exhibit a polynomial dependence on the state variable. Another important characteristic are the discontinuities of the dynamics along the boundaries of the state space, like the condition $x_i \geq y_i^-$, for $1\leq i\leq n$ in our example. This is referred to as boundary effects.

The polynomial property of the reaction rates has the consequence that multiple timescales may drive, sometimes very fast, the time evolution of CRNs. As it will be seen, in Section 7 for example, the discontinuities of the kinetics can on their side slow down significantly the time evolution of some of the components of the Markov process.

In this paper we focus on the transient behavior of these networks. We study the situation when the CRN starts from a "large" state and investigate how it goes back to a neighborhood of the origin. Depending on which component is large, the timescale to return to 0 can be very different. This can be helpful to study the property of positive recurrence but, more important perhaps, it gives insights on the time evolution of these complex Markov processes.

1.1. Lyapunov Criteria for Positive Recurrence Property of CRNs. A classical and important stability result of dynamical systems associated to a class of deterministic CRNs has been established in Feinberg [26]. It states that if a CRN is weakly reversible with deficiency 0 then there is a unique positive equilibrium point (c_i) which is locally stable. See Section 2.A of the Appendix for the definitions. The dynamical system is defined by Relation (2.15).

In a stochastic context, Anderson et al. [7] have shown that an invariant measure of the Markov process $(X(t))$ associated to a weakly reversible with deficiency 0 CRN with n chemical species exists and is expressed as a product of n Poisson distributions whose parameters are given by the coordinates of the equilibrium of the associated deterministic CRN.

Outside this class of networks, there are few general results establishing the positive recurrence of Markov processes associated to CRNs. The use of a Lyapunov function is a possible tool in this domain. If $\mathcal{G}$ is the infinitesimal generator of the Markov process, it amounts to find an *energy function*, i.e. a non-negative function f on the state space $\mathcal{S}$ converging to infinity when the norm of the state gets large, satisfying the Lyapunov condition: there exist $K\geq 0$ and $\gamma>0$, such that, for any $x\in\mathcal{S}$,

$$\mathcal{G}(f)(x) \leq -\gamma, \tag{2.1}$$

if $\|x\|\geq K$. There are several examples of specific CRNs using entropy as a Lyapunov function, see Anderson et al. [9, 10], Anderson and Kim [8], and Xu et al. [79]. Agazzi and Mattingly [4] use polynomial functions. See also Anderson et al. [6].

Constructing a Lyapunov function is, in general, a tricky problem since the expression of $\mathcal{G}(f)(x)$ involves a summation, depending on x, on, a priori, the whole state space. It can be easy to derive "local" Lyapunov functions for states x with some specific coordinates large, for example. The problem is then of having a *global* Lyapunov function, i.e. of gluing in some way all these local Lyapunov functions. The analysis of Agazzi and Mattingly [4] illustrates quite well the difficulty of constructing a Lyapunov function.

A natural extension of the criterion (2.1) simplifies significantly this problem in fact. The original version is due to Filonov [30]. It states essentially that if there exist an energy function f, an integrable stopping time τ and constants $K\geq 0$ and $\gamma>0$, such that, for any $x\in\mathcal{S}$, the relation

$$\mathbb{E}_x(f(X(\tau)))-f(x) \leq -\gamma\mathbb{E}_x(\tau), \tag{2.2}$$

holds if $\|x\|\geq K$, then $(X(t))$ is positive recurrent. See Theorem 2.3 for a precise statement.

As it can be seen, if $\tau=t_1$, t_1 being the first jump of $(X(t))$, then, by Dynkin's Formula, provided that convenient integrability properties hold, for $x\in\mathcal{S}$,

$$\mathbb{E}_x(f(X(\tau)))-f(x) = \mathbb{E}_x\left(\int_0^\tau \mathcal{G}(f)(X(s))\,\mathrm{d}s\right) = \mathcal{G}(f)(x)\mathbb{E}_x(\tau),$$

so that Condition (2.2) is in fact the classical criterion (2.1) for this choice of τ.

We can express Condition (2.1) as the fact that, starting from a large state, the energy of $(X(t))$ decreases right away, i.e. at the first jump of the process. As we will see, this property does not hold in general for many examples, especially for

some interesting CRNs. Condition (2.2) is natural in the sense that if the process, starts from a large state, it may take some time before its energy decreases. In some cases considering the instants of the second, or third jump is enough. In other example, we will have to take instants of jumps with a random index.

With Condition (2.2), roughly speaking, starting from $x \in \mathcal{S}$, the choice of an instant τ determines, i.e. with a high probability, the type of jump at time τ and therefore the decay $f(X(\tau))-f(x)$. The partitioning of the state space is on the initial state x to define the convenient stopping time τ on each subset of the partition. Some intuition may suggest the "correct" variable. Last but not least, the choice of the function f for Condition (2.2) is not a problem in general, a linear function is enough, except in Section 7.1 where a polynomial function is used. For Condition (2.1), the situation is quite different, one has to partition of the state space for the definition of f, so that the global quantity $\mathcal{G}(f)(x)$ is negative. This may be in some cases a more obscure procedure even if some intuition may also help. See Agazzi and Mattingly [4].

One of the goals of this paper is of stressing the potential of Criterion (2.2) to study positive recurrence of CRNs. It has been already used for some specific choices of τ. The general formulation (2.2) should be considered as a standard tool in this domain. All positive recurrence results of our paper are established in this way.

1.2. Scaling Analysis with the Norm of the Initial State. For the study of the *transient* behavior of CRNs, we proceed as follows: for a convenient set of coefficients $\alpha=(\alpha_i)\in\mathbb{R}_+^n$ and $\beta\in\mathbb{R}$ and if x is the initial state of $(X(t))$, it is expressed as the investigation of the convergence in distribution of the family of the sample paths

$$\left(\overline{X}_x(t)\right) \stackrel{\text{def.}}{=} \left(\frac{1}{\|x\|^{\alpha_i}} X_i\left(\frac{t}{\|x\|^{\beta}}\right), i=1,\dots,n\right),$$

when the norm of x goes to infinity. In general the large initial states "near" one of the boundaries of the CRNs are the most challenging ones as it will be seen in the examples.

The choice of β is important to identify the chemical reactions driving the time evolution on specific timescales and to establish a corresponding functional limit theorem for the scaled process.

In general there does not exist the notion of *fluid limits*, in the sense that the *global* convergence in distribution of $(\overline{X}_x(t))$ as $\|x\|$ goes to infinity does not hold in general. Several choices are in general possible for $\alpha\in\mathbb{R}_+^n$ and $\beta\geq 0$, and this is not a simple task as it will be seen. The constant β for the timescale and the coefficients (α_i) for the space variable may depend, for example, on how the sequence of initial states is going to infinity. Worse, for some sequence of initial states, there does not seem to exist any such constants (α_i) and β for a convergence in distribution result of at least a subsequence. This is where the general Lyapunov criterion is useful. For a quick discussion on fluid limits, see Section 4.4.

The main goal of this paper in this domain is of showing, via the analysis of interesting examples, how a scaling analysis of this type can give insights on the diversity of possible time evolutions of CRNs. The proofs of the associated functional limit theorems use several classical tools of probability theory: stochastic calculus, averaging principles, time change arguments, hitting time estimates, ... If they have been already used in the work of Kurtz and co-authors in a different scaling framework, their potential seems nevertheless to be generally underestimated in the literature. This is the reason for which several examples are worked out in detail in our paper.

Overview of the Paper. The formal definitions and notations are introduced in Section 2. Classical results of the literature of CRNs, including the deficiency zero theorem, are recalled in Section 2.A of the Appendix.

Lyapunov criteria for positive recurrence of Markov processes are presented and discussed in a general framework in Section 3 .

The scaling with the norm of the initial state is introduced in Section 4. The scalings used in the literature of CRNs are presented and discussed. The important feature of this scaling is that, contrary to the existing scalings, it does not change the reaction rates of the CRN or its associated graph. It is essentially focused on the time evolution of the CRNs starting from a "large" state.

The rest of the paper is devoted to the investigation of several examples of CRNs to illustrate the insight that can be obtained on the time evolution of CRNs from this scaling perspective, and, in addition, of the efficiency of the general Lyapunov criterion.

Section 5 is devoted to the analysis of some binary CRNs, that are chemical reaction networks with two chemical species and there are the complexes of reaction have at most two molecules. Two examples of *triangular* networks are considered both of them have the complexes $\{S_1, S_2, S_1+S_2\}$ and, for one of them, also a sink/source $\emptyset$. The proofs in this section are essentially standard, the main motivation is to show how Filonov's result can be useful, and also, that even in this simple setting, there are cases with three regimes corresponding to different timescales and different functional limit theorems.

For the sake of completeness, in Section 2.C of the Appendix we present some of the classical technical arguments used, repeatedly, in the proofs of convergence in distribution of sequence of processes. It includes the proof of functional laws of large numbers and of an averaging principle.

In Section 2.D of the Appendix, a general result for the stability of triangular networks with arbitrary complexes at the vertices is shown. This is an analogue of Theorem 7.6.1 of Feinberg [27] for star networks.

Section 6 is devoted to the analysis of an interesting CRN proposed in Agazzi and Mattingly [4]. The purpose of this reference is of showing that with a small modification of the graph structure of a CRN, its associated Markov process can be either positive recurrent, null recurrent, or transient. The main technical part of the paper is devoted essentially to the construction of a Lyapunov function satisfying Condition (2.1) for one graph structure. We show that Condition (2.2) can be in fact used with a simple function to prove the positive recurrence. Additionally, a scaling picture for the time evolution of this CRN is also presented.

In Section 7, a thorough analysis of the following CRN, for $p \geq 2$,

$$\emptyset \underset{\kappa_1}{\overset{\kappa_0}{\rightleftharpoons}} S_1+S_2, \qquad pS_1+S_2 \underset{\kappa_3}{\overset{\kappa_2}{\rightleftharpoons}} pS_1+2S_2,$$

is achieved. In the context of a large class of CRNs with two species analyzed in Agazzi et al. [2], this is an important example introduced and discussed from the point of view of its stability properties for $p=2$. It shows how boundary effects can complicate significantly the verification of a Lyapunov criterion (2.1) with the entropy function used as an energy function.

The analysis of this, apparently simple, CRN is in fact quite demanding. We investigate this CRN with the general approach proposed in this paper. The analysis involves a large set of technical tools. It is a good example of the complexity of the analysis of some CRNs. We show how Condition (2.2) can be used for positive recurrence and that a scaling analysis gives an interesting insight for the time evolution of this CRN.

Its transient behavior exhibits several unusual scaling properties which we describe quickly. This is the main motivation of our detailed analysis of Section 7. Such a behavior is probably not an exception in the context of CRNs. Additionally, it gives an interesting example of the use of time change arguments to derive scaling results for multi-timescales models.

The network has a kind of bi-modal behavior due to its boundary conditions. This property is exhibited via a scaling analysis of this CRN for two classes of initial states. The corresponding limiting results are:

(a) For an initial state of the form (N, b), with $b \in \mathbb{N}$ fixed.
Theorem 2.23 shows that there exists some $t_\infty > 0$, such that the convergence in distribution of processes

$$\lim_{N \to +\infty} \left(\frac{X_1(Nt)}{N}, t < t_\infty \right) = \left(1 - \frac{t}{t_\infty}, t < t_\infty \right),$$

holds with

$$t_\infty = 1 \Big/ \kappa_0 \left(e^{\kappa_2/\kappa_3} - 1 \right).$$

(b) If the initial state is of the form (a, N), $a < p$.
For the convergence in distribution of its occupation measure, see Definition 2.28, the relation

$$\lim_{N \to +\infty} \left(\frac{X_2(N^{p-1}t)}{N} \right) = (V(t))$$

holds, where $(V(t))$ is an *explosive* Markov process on $(0, 1]$ with a multiplicative structure, its infinitesimal generator $\mathcal{A}$ is given by

$$\mathcal{A}(f)(x) = \frac{r_1}{x^{p-1}} \int_0^1 \left(f\left(x u^{\delta_1} \right) - f(x) \right) \mathrm{d}u, \quad x \in (0, 1],$$

for any Borelian function f on $(0, 1)$, with $\delta_1 = \kappa_3 (p-1)!/\kappa_1$. This process is almost surely converging to 0.

In this example, to decrease the norm of the process, one has to use the timescale (Nt) in (a) and $(N^{p-1}t)$ in (b) and the decay in (a) is linear with respect to time. This is in contrast with the examples of Sections 5 and 6 where the "right" timescale is of the form (t/N^β) with $\beta \in \{0, 1/2, 1\}$. Note also that the limit of the first order of (a) is a *random* process, instead of a classical deterministic function solution of an ODE as it is usually the case.

General Remarks.

(a) Case Study.
If we are stressing the interest of using of Condition (2.2) in a general situation for CRNs. The analysis of the examples we have chosen, in Section 7 especially, are, in our view, representative of the challenging problems to derive limit theorems on the transient behaviors of these networks. There are other types of difficulties, like the interaction of more than two timescales for example, which are not represented. There is a lot of work ahead !

(b) Weakly Reversible Examples.
It should be noted that several of our examples in Sections 5 and 7 for example, are weakly reversible CRNs with deficiency 0, and thus are positive recurrent by a result of Anderson et al. [7]. See Section 2.A for the definitions. Our main point in this paper is on the interest of using convenient Lyapunov criteria and scaling ideas to investigate CRNs. We have chosen simple examples for this purpose. The fact that some of them are weakly

reversible with deficiency 0 is not really important, their simplicity is, to avoid additional technical complications essentially.

In the examples of Sections 5, if some chemical reaction are added to the CRN, then the assumption of the stability result of Anderson et al. [7] will break down, but the method used here is still valid. It should be also noted that if the example of Section 7 is positive recurrent by the deficiency 0 theorem, the scaling properties we prove for this CRN give, in our view, a quite interesting insight on the time evolution of this CRN.

2. Mathematical Models of CRNs

2.1. General Notations. For $z{=}(z_i){\in}\mathbb{N}^n$, we denote

$$\|z\| \stackrel{\text{def.}}{=} \sum_{i=1}^{n} |z_i|, \text{ and } \|z\|_\infty \stackrel{\text{def.}}{=} \max_{1\le i\le n} |z_i|. \tag{2.3}$$

We define the generalized factorials, for $z{=}(z_i){\in}\mathbb{N}^n$ and $y{=}(y_i){\in}\mathcal{C}$,

$$z! \stackrel{\text{def.}}{=} \prod_{i=1}^{n} z_i!, \quad z^{(y)} \stackrel{\text{def.}}{=} \frac{z!}{(z-y)!} = \prod_{i=1}^{n} \frac{z_i!}{(z_i-y_i)!}, \tag{2.4}$$

with the convention that $z^{(y)}{=}0$, if there exists some $i_0{\in}\mathcal{S}$ such that $y_{i_0}{>}z_{i_0}$.

2.1.1. *Markov processes.* A real-valued function $(x(t))$ on $\mathbb{N}^n$ is càdlàg if it is right continuous and it has left-limits everywhere on $\mathbb{R}_+$, for $t{>}0$, $x(t{-})$ denotes the left limit at t.

If $(X(t))$ is a càdlàg Markov process on $\mathbb{N}^n$, it is assumed that on the probability space, there is a semi-group of *shift operators* (θ_t) so that the relation $X(t{+}s){=}X(t){\circ}\theta_s$ holds almost surely for all t, $s{\ge}0$. See Chapter I of Sharpe [72] for a general presentation of this formalism for Markov processes.

We will denote by

$$t_1 = \inf\{s{>}0 : X(s) \neq X(s{-})\}, \tag{2.5}$$

the first instant of jump of $(X(t))$ and for $n{\ge}1$,

$$t_{n+1} = \inf\{s{>}t_n : X(s) \neq X(s{-})\} = t_n{+}t_1{\circ}\theta_{t_n},$$

the sequence of the instants of successive jumps of the process. The Markov process is non-explosive if and only if the sequence (t_n) is almost surely converging to infinity. There is no reference to $(X(t))$ for the notation of the sequence (t_n), for the sake of simplicity, as long as there is no ambiguity.

2.1.2. *Random measures.* If Λ is a positive Borelian measure on $\mathbb{R}_+^2$ and A is a Borelian subset of $\mathbb{R}_+$, $A{\in}\mathcal{B}(\mathbb{R}_+)$, we will use the following notation for the differential term

$$\Lambda(A, \mathrm{d}t) = \int \mathbb{1}_{\{s\in A\}} \Lambda(\mathrm{d}s, \mathrm{d}t),$$

i.e. if f is a non-negative Borelian function on $\mathbb{R}_+$,

$$\int_{\mathbb{R}_+} f(t)\Lambda(A, \mathrm{d}t) = \int_{\mathbb{R}_+^2} f(t)\mathbb{1}_{\{s\in A\}}\Lambda(\mathrm{d}s, \mathrm{d}t).$$

2.2. General Definitions for CRNs. We now give the formal definitions for chemical reaction networks.

Definition 2.1. *A* chemical reaction network *(CRN) with* n chemical species, $n{\ge}1$, *is defined by a triple* $(\mathcal{S}, \mathcal{C}, \mathcal{R})$,

— $\mathcal{S}{=}\{1, \dots, n\}$ *is the set of chemical species;*
— $\mathcal{C}$, *the set of* complexes, *is a finite subset of* $\mathbb{N}^n$*;*
— $\mathcal{R}$, *the set of* chemical reactions, *is a subset of* $\mathcal{C}^2$.

A chemical species $j\in\mathcal{S}$ is also represented as S_j. A complex $y\in\mathcal{C}$, $y=(y_j)$ is composed of y_j *molecules* of species $j\in\mathcal{S}$, its *size* is $\|y\|=y_1+\cdots+y_n$. It is also described as

$$y=\sum_{j=1}^{n} y_j S_j.$$

The state of the CRN is given by a vector $x=(x_i, 1\leq i\leq n)\in\mathbb{N}^n$, for $1\leq i\leq n$, x_i is the number of copies of chemical species S_i. A chemical reaction $r=(y_r^-, y_r^+)\in\mathcal{R}$ corresponds to the change of state, for $x=(x_i)$,

$$x \longrightarrow x+y_r^+-y_r^- = \left(x_i+y_{r,i}^+-y_{r,i}^-, 1\leq i\leq n\right) \tag{2.6}$$

provided that $y_{r,i}^-\leq x_i$ holds for $1\leq i\leq n$, i.e. there are at least $y_{r,i}^-$ copies of chemical species of type i, for all $i\in\mathcal{S}$, otherwise the reaction cannot happen. We define

$$y_{\max}^{\pm}=\max_{r\in\mathcal{R}}\|y_r^{\pm}\|. \tag{2.7}$$

Such a chemical reaction is classically represented as

$$\sum_{i=1}^{n} y_{r,i}^- S_i \rightharpoonup \sum_{i=1}^{n} y_{r,i}^+ S_i,$$

A CRN can be seen as an oriented graph, called the *reaction graph*, whose vertices are complexes, i.e. in $\mathcal{C}$ and whose set of directed edges is $\mathcal{R}$.

The notation $\emptyset$ refers to the complex associated to the null vector of $\mathbb{N}^n$, $\emptyset=(0)$. For $y=(y_i)\in\mathcal{C}$, a chemical reaction of the type $(\emptyset, y)$ represents an external source creating y_i copies of species i, for $i=1,\ldots, n$. A chemical reaction of the type $(y, \emptyset)$ consists in removing y_i copies of species i, for $i=1,\ldots, n$, provided that there are sufficiently many copies of each species.

2.3. Markov Processes: Law of Mass Action. A stochastic model of a CRN is represented by a continuous time Markov jump process $(X(t))=(X_i(t), i=1,\ldots,n)$ with values in $\mathbb{N}^n$. The dynamical behavior of a CRN, i.e. the time evolution of the number of copies of each of the n chemical species is governed by *the law of mass action*. See Voit et al. [76], Lund [56] for surveys on the law of mass action.

For these kinetics, the associated Q-matrix of $(X(t))$ is defined so that, for $x\in\mathbb{N}^n$ and $r=(y_r^-, y_r^+)\in\mathcal{R}$, the transition $x\to x+y_r^+-y_r^-$ occurs at rate

$$\kappa_r x^{(y_r^-)} = \kappa_r \frac{x!}{(x-y_r^-)!} = \kappa_r \prod_{i=1}^{n} x_i(x_i-1)\cdots(x_i-y_{r,i}+1), \tag{2.8}$$

where $\kappa=(\kappa_r, r\in\mathcal{R})$ is a vector of non-negative numbers, for $r\in\mathcal{R}$, κ_r is the reaction rate of r. More formally the functional operator $\mathcal{Q}(f)$ associated to its Q-matrix is defined by, for $x\in\mathbb{N}^n$,

$$\mathcal{Q}(f)(x)=\sum_{r\in\mathcal{R}} \kappa_r x^{(y_r^-)} \left(f\left(x+y_r^+-y_r^-\right)-f(x)\right), \tag{2.9}$$

for any function f with finite support on $\mathbb{N}^n$.

2.4. A SDE Formulation. The Markov process with Q-matrix defined by Relation (2.8) can be classically expressed as the solution of a martingale problem. See Theorem (20.6) in Section IV of Rogers and Williams [69].

We assume that on the probability space we have a set of independent Poisson point processes $\mathcal{P}_r$, $r\in\mathcal{R}$ on $\mathbb{R}_+^2$ with intensity measure the Lebesgue measure on

$\mathbb{R}_+^2$. See Kingman [64]. The Markov process has the same distribution as the solution $(X(t))=(X_i(t))$ of the SDE,

$$\mathrm{d}X(t) = \sum_{r=(y_r^-,y_r^+)\in\mathcal{R}} \left(y_r^+ - y_r^-\right) \mathcal{P}_r\left(\left(0, \kappa_r \frac{X(t-)!}{(X(t-)-y_r^-)!}\right), \mathrm{d}t\right). \tag{2.10}$$

Note that a solution of SDE (2.10) is not, a priori, defined on the entire half-line since the instants of jumps of the process may converge to some random variable T_∞, the time of explosion. When this variable is finite, a point † is added to the state space, with the convention $X(t)=\dagger$ for $t\geq T_\infty$.

The associated filtration is $(\mathcal{F}_t)$, where, for $t\geq 0$, $\mathcal{F}_t$ is the completed σ-field generated by the random variables

$$\mathcal{F}_t \stackrel{\text{def.}}{=} \sigma\left(\mathcal{P}_r(A\times[0,s)), r\in\mathcal{R}, s\leq t, A\in\mathcal{B}(\mathbb{R}_+)\right). \tag{2.11}$$

The notions of stopping times and martingales refer implicitly to this filtration.

Provided that $(X(t))$ is well defined on $[0,T]$, $T>0$, the integration of SDE (2.10) gives the relation

$$X(t) = X(0) + \sum_{r\in\mathcal{R}} M_r(t) + \sum_{r\in\mathcal{R}} \kappa_r \left(y_r^+ - y_r^-\right) \int_0^t \frac{X(s)!}{(X(s)-y_r^-)!}\,\mathrm{d}s \tag{2.12}$$

on the time interval $[0,T]$, where, for $r\in\mathcal{R}$, $(M_r(t))=(M_{r,i}(t))$ is a local martingale defined by

$$\left(\left(y_r^+ - y_r^-\right) \int_0^t \left(\mathcal{P}_r\left(\left(0, \kappa_r \frac{X(s-)!}{(X(s-)-y_r^-)!}\right), \mathrm{d}s\right) - \kappa_r \frac{X(s)!}{(X(s)-y_r^-)!}\,\mathrm{d}s\right)\right), \tag{2.13}$$

and its previsible increasing process is given by, for $1\leq i,j\leq n$,

$$\left(\langle M_{r,i}, M_{r,j}\rangle(t)\right) = \left(\left(y_{r,i}^+ - y_{r,i}^-\right)\left(y_{r,j}^+ - y_{r,j}^-\right)\kappa_r \int_0^t \frac{X(s)!}{(X(s)-y_r^-)!}\,\mathrm{d}s\right). \tag{2.14}$$

For $q\in\mathcal{R}$, $q\neq r$, $\left(\langle M_{r,i}, M_{q,j}\rangle(t)\right)\equiv 0$.

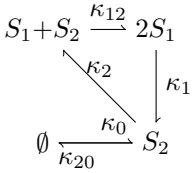


FIGURE 1. Example of a CRN

Example. For the CRN of Figure 1, we have

- $\mathcal{S}=\{1,2\}$, $\mathcal{C}=\{\emptyset, S_1+S_2, 2S_1, S_2\}$;
- $\mathcal{R}=\{(\emptyset,S_2),(S_2,\emptyset),(S_2,S_1+S_2),(S_1+S_2,2S_1),(2S_1,S_2)\}$.

The non-trivial components of the Q-matrix of the associated Markov process $(X(t))=(X_1(t),X_2(t))$ are given by

$$(x_1,x_2) \longrightarrow (x_1,x_2) + \begin{cases} (0,1), & \kappa_0, \\ (0,-1), & \kappa_{20}x_2, \end{cases} \qquad (x_1,x_2) + \begin{cases} (1,0), & \kappa_2\, x_2, \\ (-2,1), & \kappa_1\, x_1(x_1-1), \\ (1,-1), & \kappa_{12}\, x_1 x_2. \end{cases}$$

The associated SDEs for $(X_1(t), X_2(t))$ are

$$\begin{cases} \mathrm{d}X_1(t) = \mathcal{P}_2((0,\kappa_2 X_2(t-)),\mathrm{d}t)+\mathcal{P}_{12}((0,\kappa_{12}X_1X_2(t-)),\mathrm{d}t) \\ \qquad\quad -2\mathcal{P}_1((0,\kappa_1 X_1(t-)(X_1(t-)-1)),\mathrm{d}t) \\ \mathrm{d}X_2(t) = \mathcal{P}_1((0,\kappa_1 X_1(t-)(X_1(t-)-1)),\mathrm{d}t)+\mathcal{P}_0((0,\kappa_0),\mathrm{d}t) \\ +\mathcal{P}_1((0,\kappa_1 X_1(t-)),\mathrm{d}t)-\mathcal{P}_{12}((0,\kappa_{12}X_1X_2(t-)),\mathrm{d}t)-\mathcal{P}_{20}((0,\kappa_{20}X_2(t-)),\mathrm{d}t). \end{cases}$$

2.5. An Important CRN: The M/M/∞ queue. This is a simple CRN with an external input and one chemical species,

$$\emptyset \underset{\mu}{\overset{\lambda}{\rightleftharpoons}} S_1.$$

The $M/M/\infty$ queue with input parameter $\lambda\geq 0$ and output parameter $\mu>0$ is a Markov process $(L(t))$ on $\mathbb{N}$ with transition rates

$$x \longrightarrow \begin{cases} x+1 & \lambda \\ x-1 & \mu x. \end{cases}$$

The invariant distribution of $(L(t))$ is Poisson with parameter $\rho=\lambda/\mu$.

This fundamental process can be seen as a kind of discrete Ornstein-Uhlenbeck process. It has a long history, it has been used in some early mathematical models of telephone networks at the beginning of the twentieth century, see Erlang [23], also in stochastic models of natural radioactivity in the 1950's, see Hammersley [38] and it is the basic process of mathematical models communication networks analyzed in the 1970's, see Kelly [61]. See Chapter 6 of Robert [67].

Technical results on this stochastic process turn out to be useful to investigate the scaling properties of some CRNs and, as we will see, in the construction of couplings used in our proofs. See Sections 6 and 2.C.1.2 for example.

2.6. Deterministic Models of CRNs. Early mathematical models of CRNs are continuous models with state space $\mathbb{R}^n_+$. They have been used in chemical physics to describe the time evolution of *concentrations* of the n chemical species in a large volume, instead of the vector of the number of copies. The process is a deterministic dynamical system in $\mathbb{R}^n_+$, $(x(t))=(x_i(t))$ expressed as the solution of the following set of Ordinary Differential Equations (ODEs),

$$\dot{x}(t) = \sum_{r=(y_r^-,y_r^+)\in\mathcal{R}} \kappa_r \left(\prod_{i=1}^n x_i(t)^{y_{r,i}^-}\right)\left(y_r^+ - y_r^-\right), \tag{2.15}$$

with also a polynomial dependence of the state variable. This is the original *law of mass action* in fact, see the historical reference Guldberg and Waage [35]. See Relation 2.15. It should be noted that, due to a possible lack of a global Lipschitz property, a solution of ODE (2.15) may also blow-up, and be defined on a finite time interval only.

A striking and impressive result in this context is the *Deficiency Zero Theorem*. See Section 2.A of the Appendix.

3. Filonov's Stability Criterion

In this section we formulate a criterion, due to Filonov [30], of positive recurrence for continuous time Markov processes associated to CRNs. It is an extension of the classical Lyapounov criterion, see Corollary 2.4. Our experience is that it is really useful in the context of CRNs and more practical than the classical criterion.

Definition 2.2. *An energy function f on $\mathcal{E}_0$ is a non-negative function such that, for all $K>0$, the set $\{x\in\mathcal{E}_0 : f(x)\leq K\}$ is finite.*

By convention the value of an energy function at the point at infinity † is $+\infty$. The classical energy functions used in a CRN framework can be linear functions or an entropy function, for $x=(x_i)$,

$$f(x)=\sum_{i=1}^{m} a_i x_i, \quad f(x)=\sum_{i=1}^{m}(x_i \ln x_i - x_i + 1).$$

THEOREM 2.3 (Filonov). *Let $(X(t))$ be an irreducible Markov process on $\mathcal{E}_0 \subset \mathbb{N}^m$ associated to a CRN network with Q-matrix (2.8), if there exist*

(a) *an integrable stopping time τ such that $\tau \geq t_1 \wedge \eta$, for a constant $\eta>0$ and t_1 is the first jump of $(X(t))$, Relation (2.5);*

(b) *an energy function f on $\mathcal{E}_0$ and constants K and $\gamma>0$ such that the relation*

$$\mathbb{E}_x\left(f(X(\tau))\right) - f(x) \leq -\gamma \mathbb{E}_x(\tau), \tag{2.16}$$

holds for all $x \in \mathcal{E}_0$ such that $f(x) \geq K$,

then $(X(t))$ is a positive recurrent Markov process.

A function f satisfying Condition (2.16) is usually referred to as a *Lyapunov function*.

A difficulty of the continuous time case is the possibility of explosion, when the number of jumps if infinite in a finite time interval. It may be seen pathological, as we will see, due to the kinetics of the law of mass action, it may happen in the context of CRNs. If there is explosion, the Markov process is of course transient. Note that if Equation (2.16) holds, $\tau < T_\infty$ with T_∞ the time of explosion (possibly equal to $+\infty$). A consequence of the theorem is that the process cannot explode under its assumptions.

The following positive recurrence criterion involving the Q-matrix Q of the Markov process is quite often used for the stability results of CRNs.

COROLLARY 2.4. *Let $(X(t))$ be an irreducible Markov process on $\mathcal{E}_0$ associated to a CRN network with Q-matrix Q defined by Relation (2.8), if there exists an energy function f on $\mathcal{S}$ and γ, $K>0$ such that*

$$Q(f)(x) \leq -\gamma, \text{ if } f(x) \geq K, \tag{2.17}$$

then $(X(t))$ is positive recurrent.

PROOF. For $\tau = t_1$, and x such that $f(x) \geq K$, we have

$$\mathbb{E}_x\left(f(X(t_1)) - f(x)\right) = \mathbb{E}_x\left(\int_0^{t_1} Q(f)(X(s))\,\mathrm{d}s\right) = Q(f)(x)\mathbb{E}_x(t_1) \leq -\gamma \mathbb{E}_x(t_1),$$

then we conclude with Theorem 2.3. □

The classical Condition (2.17) states that if the initial state has a large level of energy, the next step decreases in average the energy by some fixed positive quantity. There are various versions of this type of result. One of them evaluates the expected value of $f(X(t_0))$ at a deterministic function $t_0(x)$ of the initial state, see Proposition 4.5 of [17]. Finding a global Lyapunov function may be cumbersome in general. It may imply to partition the state space, for the location of $X(t_0)$, to build and to glue piecewise, local, Lyapunov functions. See Agazzi and Mattingly [4] for a typical example. For this theorem, one has to partition the state space from the point of view of the initial state, to define τ, instead of the state at time t_0. This is in general much simpler to handle.

Starting from a state x of high energy does not necessarily lead quickly to a lower energy level, i.e. $Q(f)(x)$ is not necessarily negative. Instead, it may happen that one may have to wait for some amount of time, the quantity τ, before the energy can significantly decrease. The "price" to pay for waiting is that the decay

needed is not anymore some fixed negative quantity, but, as it can be expected, a negative constant proportional to $\mathbb{E}_x(\tau)$, the mean value of the waiting time.

For the somewhat specific condition $\tau \geq t_1 \wedge \eta$, it should be noted that, at least some condition is required, since the trivial $\tau \equiv 0$ would work otherwise. In practice, it is quite natural that the "convenient" stopping time τ should be greater than t_1, the first instant when there is a change of state for the Markov process. There are, however, situations when it is enough to consider a deterministic stopping time τ, when scaling limits of sample paths are used. Some of our examples below exhibit this feature.

(a) If $\tau \equiv \eta > 0$.
Condition (2.16) is just the classical *Foster-Lyapunov condition* for the discrete Markov chain $(M_n)=(X(n\eta))$, i.e. the process $(X(t))$ on the discrete timescale $n \mapsto n\eta$,

$$\mathbb{E}_x(f(M_1)) - f(x) \leq -\gamma\eta, \quad \text{for } x \in \mathcal{S}_0 \text{ such that } f(x) > K. \tag{2.18}$$

See Hairer [37], Bramson [17], and Meyn and Tweedie [58].

(b) If $\tau = t_1$.
Condition (2.16) is equivalent to condition (2.17), see the proof of corollary 2.4. Most of stability analyses of CRN networks use this condition.

Proof of Theorem 2.3. The proof uses essentially the same arguments as in the proof of Theorem 8.6 of [67]. The only specific difficulty lies in the fact that, a priori, the Markov process can be explosive. Define the sequence of induced stopping times (s_n) by induction, by $s_1 = \tau$ and

$$s_{n+1} = s_n + \tau \circ \theta_{s_n}. \tag{2.19}$$

By using the strong Markov property of $(X(t))$, see Theorem (9.4), Section III.9 of Rogers and Williams [68], (s_n) is a non-decreasing sequence of stopping times. From our assumption on τ, we get that (s_n) is almost surely an increasing sequence, i.e. $s_n < s_{n+1}$ for all $n \geq 1$.

We define

$$T_K = \inf\{s \geq 0 : f(X(s)) \leq K\}, \text{ and } \nu \stackrel{\text{def.}}{=} \inf\{n \geq 0 : f(X(s_n)) \leq K\},$$

then, clearly $T_K \leq s_\nu$. Let, for $n \geq 1$,

$$Z_n \stackrel{\text{def.}}{=} f(X(s_n)) + \gamma s_n,$$

then, with Relation (2.16) and the strong Markov property of $(X(t))$ for the stopping time s_n, we obtain the relation

$$\mathbb{E}\left(Z_{n+1} \mid \mathcal{F}_{s_n}\right) = Z_n + \mathbb{E}_{X(s_n)}\left(f(X(\tau)) + \gamma\tau\right) - f(X(s_n)) \leq Z_n, \tag{2.20}$$

on the event $\{\nu > n\}$.

The process $(Z_{n \wedge \nu})$ is therefore a non-negative super-martingale, in particular it is converging almost surely to a finite limit.

First, lets show that ν is almost surely finite. The almost sure convergence of $(Z_{n \wedge \nu})$ gives that, almost surely on the event $\{\nu = +\infty\}$, the sequence (s_n) is converging to a finite limit. In particular the increments $(s_{n+1} - s_n)$ are less than $\eta/2$ after some finite index n_0. Our assumption, $\tau \geq t_1 \wedge \eta$ implies therefore that, for $n > n_0$,

$$s_{n+1} \geq s_n + t_1 \circ \theta_{s_n}, \tag{2.21}$$

and, by induction, $s_n \geq t_{n-n_0} \circ \theta_{s_{n_0}}$ holds for $n > n_0$.

Since $X(s_n)=X(u_n)$, where $u_n \leq s_n$ is the last jump instant of $(X(t))$ before s_n, we have $u_n < u_{n+1}$ by Relation (2.21). Hence, almost surely, on the event $\{\nu=+\infty\}$, the Markov process $(X(t))$ explodes in finite time, so that

$$\limsup_{n\to+\infty} f(X(u_n)) = \limsup_{n\to+\infty} f(X(s_n)) = +\infty.$$

Since $(Z_{n\wedge\nu})$ is converging almost surely to a finite limit, this implies that the random variable ν is almost surely finite.

By integrating Relation (2.20) we get that

$$\gamma \mathbb{E}_x(s_{n\wedge\nu}) \leq \mathbb{E}_x\left(Z_{n\wedge\nu}\right) \leq \mathbb{E}_x(Z_0) = f(x),$$

the monotone convergence theorem gives the relation

$$\mathbb{E}(T_K) \leq \mathbb{E}(s_\nu) \leq \frac{f(x)}{\gamma},$$

we conclude with Proposition 2.32 of the appendix. The theorem is proved. □

4. Scaling Methods

The scaling approaches presented in this section aim at providing a first order description of the time evolution of a CRN, to investigate its transient behavior in particular.

4.1. Deterministic CRNs as a Scaled Version of Stochastic CRNs. It is possible to see the set of ODEs (2.15) with the parameters (κ_r) as a first order approximation of a scaled stochastic model of a CRN with the same triple $(\mathcal{S},\mathcal{C},\mathcal{R})$ but with a set of scaled reaction rates

$$\left(\kappa_r^N, r\in\mathcal{R}\right) = \left(\frac{\kappa_r}{N^{\|y_r^-\|-1}}, r\in\mathcal{R}\right). \tag{2.22}$$

We denote by $(X^N(t))=(X_i^N(t))$ the associated Markov jump process.

This ad-hoc scaling is not homogeneous in the sense that it does not correspond to a uniform change of timescale. The timescale of reaction $r\in\mathcal{R}$ is "slowed down" by a factor $1/N^{\|y_r^-\|-1}$. In this manner, provided that all its coordinates are of the order of N, the rate of any chemical reaction is of the same order, $O(1)$ with respect to N. Consequently, this scaling removes an important feature of CRNs, multi-timescales, i.e. that some reactions may be change a subset of coordinates of $(X(t))$ on a much faster timescale.

The following proposition is a well-known result of the CRN literature for the scaling assumption of Relation (2.22). See Mozgunov et al. [60] for example.

Proposition 2.5. *If $(X^N(t))$ is the Markov process with Q-matrix given by Relation* (2.8) *with the reaction rates given by Relations* (2.22) *and if the sequence of initial states is such that*

$$\lim_{N\to+\infty}\left(\frac{X_i^N(0)}{N}\right) = x_0=(x_{0,i})\in\mathbb{R}_+^m,$$

then the convergence in distribution

$$\lim_{N\to+\infty}\left(\frac{X_i^N(t)}{N}, t\in(0,t_\infty)\right) = (x_i(t), t\in(0,t_\infty)),$$

holds, where $(x_i(t))$ is the solution of the set of ODEs (2.15) *starting at (x_0) and t_∞ is its blow-up instant,*

$$t_\infty \overset{\text{def.}}{=} \lim_{K\to+\infty} \inf\{t{:}\|x(t)\|_\infty \geq K\},$$

with the convention that $\inf\emptyset=+\infty$.

PROOF. We give a quick proof of this classical result. For $K \geq 1$, define

$$H_K^N = \inf\left\{t>0 : \|X^N(t)\|_\infty \geq KN\right\} \text{ and } t_K = \inf\left\{t>0 : \|x(t)\|_\infty \geq K\right\},$$

and

$$\left(\overline{X}_K^N(t)\right) \stackrel{\text{def.}}{=} \left(\frac{X_i^N(t\wedge H_K^N)}{N}\right).$$

On the time interval $(0, H_K^N)$, there are at least $C_K N$ jumps of the process $(X_N(t))$, with $C_K = \lfloor (K - \|x_0\|)/y_{\max}^+ \rfloor$. See Definition (2.7). The maximal jump rate of the process on $(0, H_K^N)$ is bounded by $\overline{\lambda} N$ with $\overline{\lambda} = \|\kappa\|_\infty K^{y_{\max}^-} |\mathcal{R}|$. Hence,

$$\lim_{N\to+\infty} \mathbb{P}\left(H_K^N \leq t\right) = 0,$$

holds for any $t < C_K/\overline{\lambda}$.

We now use the SDE formulation of Section 2.4 of the appendix to represent the Markov process $(X^N(t))$. Relation (2.10) gives the identity

$$\overline{X}^N(t) = \overline{X}^N(0) + \sum_{r\in\mathcal{R}} \frac{M_r(t\wedge H_K^N)}{N} + \kappa_r \left(y_r^+ - y_r^-\right) \int_0^{t\wedge H_K^N} \frac{X^N(s)!}{N^{\|y_r^-\|}(X^N(s) - y_r^-)!}\, \mathrm{d}s, \tag{2.23}$$

where $(M_r(t)) = ((M_{r,i}(t)))$, $r \in \mathcal{R}$, is the set of martingales defined by Relation (2.13). Relation (2.14) gives, for $1 \leq i \leq m$,

$$\begin{aligned}\left\langle \sum_{r\in\mathcal{R}} \frac{M_{r,i}}{N} \right\rangle (t\wedge H_K^N) &= \sum_{r\in\mathcal{R}} \frac{\langle M_{r,i}\rangle}{N^2}(t\wedge H_K^N) \\ &= \frac{1}{N}\sum_{r\in\mathcal{R}} \left(y_{r,i}^+ - y_{r,i}^-\right)^2 \kappa_r \int_0^{t\wedge H_K^N} \frac{X(s)!}{N^{\|y_r^-\|}(X(s) - y_r^-)!}\, \mathrm{d}s \\ &\leq \frac{1}{N}\sum_{r\in\mathcal{R}} \left(y_{r,i}^+ - y_{r,i}^-\right)^2 \kappa_r K^{\|y_r^-\|} t,\end{aligned}$$

with Doob's Inequality, we get that $(M_{r,i}(t\wedge H_K^N)/N)$ is converging in distribution to 0, for all $1 \leq i \leq m$. By using the criterion of modulus of continuity, see Billingsley [14], it is easy to show that the sequence of processes $(\overline{X}_K^N(t))$ is tight for the convergence in distribution.

For $t < t_K$, with high probability we have $H_K^N \geq t$, hence by using Relation (2.23), we obtain that any limiting point of $(\overline{X}_K^N(t), t < t_K)$ satisfies the set of ODEs (2.15). In particular the sequence $(\|X_N(t)\|, t < t_K)$ is converging in distribution to the process $(\|x(t)\|, t < t_K)$. The proposition is proved. □

4.2. Alternative Scalings of Reaction Rates. Kurtz and co-authors have also investigated several examples of CRNs with analogous scaling methods. In this approach, some reaction rates may be sped-up with some power of the scaling parameter N and the state variables are scaled accordingly. The difference with the classical scaling is that the transition rates do not have necessarily the same order of magnitude in N. The initial motivation was of fitting the parameters of these scaling models with biological data obtained from experiments. Several examples of CRNs are investigated in this setting, the limit theorems derived depend, of course, on the scaling in N chosen for the vector of reaction rates (κ_r). See for example Ball et al. [11], Kang and Kurtz [46], and Kim et al. [50].

4.3. Scaling with the Norm of the Initial State. The main scaling approach considered in this paper is with respect to the size of the initial state. The topology and the vector κ of reaction rates are unchanged. With a convenient change of timescales it may give a first order description of the time evolution of the CRN starting from a "large" state.

This scaling is in fact related to Theorem 2.3, we start by a reformulation of it.

PROPOSITION 2.6. *Let $(X(t))$ be a Markov process associated to a CRN network with Q-matrix defined by Relation (2.8), if there are an energy function f on $\mathcal{E}_0$, constants ε and $\eta>0$, and an integrable stopping time $\tau\geq t_1\wedge\eta$ such that the relations*

$$\limsup_{\substack{x\in\mathcal{E}_0\\ f(x)\to+\infty}} \frac{\mathbb{E}_x(f(X(\tau)))}{f(x)} \leq 1-\varepsilon \text{ and } C_0 \stackrel{\text{def.}}{=} \limsup_{\substack{x\in\mathcal{E}_0\\ f(x)\to+\infty}} \frac{\mathbb{E}_x(\tau)}{f(x)} < +\infty$$

hold, then the Markov process $(X(t))$ is positive recurrent.

PROOF. One can find $K>0$ such that if $x\in\mathcal{E}_0$ satisfies $f(x)\geq K$ then

$$\mathbb{E}_x(f(X(\tau)))-f(x) \leq -\frac{\varepsilon}{2}f(x) \leq -\frac{\varepsilon}{4C_0}\mathbb{E}_x(\tau),$$

Theorem 2.3 concludes the proof. □

Starting from a large x, we can interpret τ in Proposition 2.6 as the time "back" in direction of the origin. If we take f as the norm $\|\cdot\|$, the above proposition suggests the introduction of the scaled processes, for $x\in\mathcal{E}_0$,

$$\left(\overline{X}_x(t)\right) \stackrel{\text{def.}}{=} \left(\frac{X\left(tg(x)\right)}{\|x\|}, t\geq 0\right),$$

where g is a positive function on $\mathcal{S}$ such that $g(x)\leq C_0\|x\|$, $x\in\mathcal{S}$, for some $C_0>0$. The function g is deterministic function related to the order of magnitude of τ starting from x.

Assume that, as $\|x\|$ goes to infinity, the family of processes $(\overline{X}_x(t))$ converges in distribution to some deterministic process $(\overline{x}(t))$. The dynamical system $(\overline{x}(t))$, with an initial point on the unit sphere of $\mathbb{R}_+^m$ provides a first order description of the Markov process. Such limiting result may be complicated to obtain, because of polynomial reaction rates and of boundary behaviors already mentioned. There are in general different timescales involved and, for this reason, there is rarely "one" function g for example. Still, it gives a method to investigate the qualitative behavior of these complicated Markov processes.

There is an interesting, global, scaling of this kind. Assuming that $\|x\|$, the norm of the initial state is large, the rate of the fastest reaction is at most of the order of $\|x\|^{y_{\max}^-}$, $y_{\max}^-$ is the size of the largest complex initiating a reaction, see Relation (2.7). For this reason, when looking at the scaled process, the timescale $t\mapsto t/\|x\|^{y_{\max}^- -1}$ is quite natural.

PROPOSITION 2.7. *If $(X^N(t))$ is the Markov process associated to a CRN network $(\mathcal{S},\mathcal{C},\mathcal{R})$ with parameters (κ_r) and whose initial state is such that*

$$\lim_{N\to+\infty}\left(\frac{X_i^N(0)}{N}\right) = x_0=(x_{0,i})\in\mathbb{R}_+^m,$$

then there exists $t_\infty>0$, such that when N goes to infinity the family of processes

$$\left(\overline{X}^N(t)\right) \stackrel{\text{def.}}{=} \left(\frac{1}{N}X\left(t\Big/N^{y_{\max}^- -1}\right), t<t_\infty\right),$$

is converging in distribution to $(\ell(t), t<t_\infty)$*, the solution of the ODE*

$$\mathrm{d}\ell(t) = \sum_{r\in\mathcal{R}, \|y_r^-\|=y^-_{\max}} \kappa_r \ell(t)^{y_r^-} (y_r^+ - y_r^-) \tag{2.24}$$

on $(0, t_\infty)$*, with* $\ell(0) = x_0$*.*

This scaling has the effect of considering the same CRN but with a set of reactions reduced to $\{r=(y_r^-, y_r^+)\in\mathcal{R}: \|y_r^-\|=y^-_{\max}\}$.

Proof. We proceed in an analogous way as in the proof of Proposition 2.5. The SDE formulation of Section 2.4 of the appendix to represent the Markov process $(X^N(t))$ is used. Relation (2.10) gives for $t\geq 0$,

$$\begin{aligned} \overline{X}^N(t) = \overline{X}^N(0) &+ \sum_{r\in\mathcal{R}} \overline{M}_r^N(t) \\ &+ \sum_{r\in\mathcal{R}} \kappa_r \left(y_r^+ - y_r^-\right) \int_0^t \frac{X^N(s/N^{y^-_{\max}-1})!}{N^{y^-_{\max}} (X^N(s/N^{y^-_{\max}-1}) - y_r^-)!}\, \mathrm{d}s, \end{aligned} \tag{2.25}$$

where, for $r\in\mathcal{R}$, $(\overline{M}_r^N(t))$ is a martingale. Stopping the process at H_K^N as in the proof of Proposition 2.5, we remark that for $r=(y_r^-, y_r^+)\in\mathcal{R}$ such that $\|y_r^-\|<y^-_{\max}$,

$$\frac{X^N(t)!}{N^{y^-_{\max}} (X^N(t) - y_r^-)!} \leq \frac{K^{y^-_{\max}-1}}{N},$$

so that the corresponding terms in Relation (2.25) vanish, as processes, when N gets large. We can show easily that taking t_∞ small enough,

$$\lim_{N\to+\infty} \mathbb{P}\left(H_K^N \leq t\right) = 0,$$

which allows us to conclude. □

4.4. Analogies with Queueing Networks. This context is reminiscent of an analogous situation for queueing networks at the beginning of the 1990's. A queueing network can be described as a finite set of processing units among which several flows of requests are circulating. A request requires the processing capacity of a subset of nodes for some amount of time before leaving the network. An important difference with CRN is that the jump rates are depending at most linearly on the state and quite often they are in fact bounded, for Jackson networks for example. Several large classes of queueing networks having product form invariant distributions were identified, like Jackson networks or loss networks, see Kelly [48]. Outside these classes, little is known, even on the existence of equilibrium distributions.

In the early 1990s, several simple queueing networks were shown to be, surprisingly, unstable. See, for example, the pioneering papers Rybko and Stolyar [70] and Bramson [16]. The limits of scaled models when the norm of the initial state is getting large, sometimes referred to as *fluid limits*, have been used for example in the mathematical analysis of Rybko and Stolyar [70]. See Bramson [17] and Chapter 9 of Robert [67], and references therein, for a general presentation. Scaling ideas have mainly been used to prove positive recurrence properties rather than investigating a first order description of these networks. Chen and Mandelbaum [66] for Jackson networks is an exception in this domain. The Markov processes associated to these queueing networks can be described as behaving locally as random walks in several subsets of $\mathbb{N}^m$. For this reason, there is essentially one relevant time scale $t\to Nt$, with $N=\|X(0)\|$, contrary to CRNs where there may be several, very different, timescales.

Because of the diversity of relevant timescales in general, a stability analysis using fluid limits, as it can be done for some stochastic models of queueing networks, seems to have a limited interest in practice in the context of CRNs.

It should be noted that for a class of queueing networks, loss networks, a different type of scaling is of interest. It consists in speeding-up the rates of external arrivals by a (large) factor N, and to study the asymptotic behavior of the process scaled by $1/N$ on the normal timescale. The motivation of this scaling is of having a first order description of the time evolution of the network. The stability property is clear, the associated Markov process being reversible. See Kelly [61].

5. Binary CRN Networks

In this section, we investigate simple examples of CRNs with complexes whose size is at most 2.

Definition 2.8. *A CRN network $(\mathcal{S},\mathcal{C},\mathcal{R})$ with n chemical species is* binary *if any complex $y\in\mathcal{C}$ is composed of at most two molecules, i.e. $\|y\|\leq 2$.*

The set of complexes can be represented as $\mathcal{C}=J_0\cup J_1\cup J_2$, where, for $i\in\{0,1,2\}$, the subset J_i is the set of complexes of size i, note that J_i can be empty. If $y\in J_1$, with a slight abuse of notation, it will be represented as $y=S_y$ for some $S_y\in\mathcal{S}$. Similarly, $y=S_y^1+S_y^2$, for $S_y^1,S_y^2\in\mathcal{S}$ when $y\in J_2$.

Proposition 2.9. *If $(X(t))$ is the Markov process associated to a binary CRN network then, almost surely, it is not explosive. Furthermore, for $T>0$, the family of random variables in $\mathbb{R}_+$*

$$(X_x^*(T), x\in\mathbb{N}^n\setminus\{0\}) \stackrel{\text{def.}}{=} \left(\frac{1}{\|x\|}\sup_{s\leq T}\|X(s)\|, X(0)=x\in\mathbb{N}^n\setminus\{0\}\right)$$

is uniformly integrable.

Proof. The SDEs associated to the CRN are given by

$$\mathrm{d}X(t)=\sum_{r\in\mathcal{R}}(y_r^+-y_r^-)\mathcal{P}_r\left(\left(0,\kappa_r\frac{X_r!}{(X_r-y_r^-)!}\right),\mathrm{d}t\right),$$

with our convention that $\mathcal{P}_r$, $r\in\mathcal{R}$, are independent Poisson processes on $\mathbb{R}_+^2$ with intensity measure $\mathrm{d}s\otimes\mathrm{d}t$ on $\mathbb{R}_+^2$. See Section 2.4 of the appendix. The binary condition implies that for $r\in\mathcal{R}$ with $y_r^-\in J_2$, then

$$\|y_r^+\|-\|y_r^-\|\leq 0.$$

If $|\mathcal{R}|$ denotes the cardinality of $\mathcal{R}$, it is not difficult to construct a coupling so that $\|X(t)\|\leq Z(t)$ holds, where $(Z(t))$ is the solution of the SDE

$$\mathrm{d}Z(t)=2\mathcal{P}\left((0,\|\kappa\|_\infty|\mathcal{R}|Z(t-)),\mathrm{d}t\right),$$

with $Z(0)=\|X(0)\|=\|x\|$. The process $(Z(t))$ is simply a pure birth branching process with rate $\|\kappa\|_\infty|\mathcal{R}|$. It is almost surely non-explosive. For $0\leq t\leq T$, we have

$$Z(t)=\|x\|+M_Z(t)+2\|\kappa\|_\infty|\mathcal{R}|\int_0^t Z(s)\,\mathrm{d}s,$$

where $(M_Z(t))$ is a martingale, with

$$\langle M_Z\rangle(t)=4\|\kappa\|_\infty|\mathcal{R}|\int_0^t Z(s)\,\mathrm{d}s.$$

It is easily seen that $\mathbb{E}(Z(t))=\|x\|\exp(2\|\kappa\|_\infty|\mathcal{R}|t)$. If

$$Z_x^*(t)\stackrel{\text{def.}}{=}\sup_{s\leq t}\left(\frac{Z(s)}{\|x\|}\right),$$

then,

$$\frac{1}{9}Z_x^*(t)^2 \leq 1+\frac{1}{\|x\|^2}\sup_{s\leq T}|M_Z(s)|^2+4T(\|\kappa\|_\infty|\mathcal{R}|)^2\int_0^t Z_x^*(s)^2\,\mathrm{d}s.$$

Doob's Inequality gives the inequality

$$\mathbb{E}\left(\sup_{s\leq T}|M_Z(s)|^2\right) \leq 4\mathbb{E}(M_Z(T)^2)$$
$$= 16\|\kappa\|_\infty|\mathcal{R}|\int_0^T \mathbb{E}(Z(s))\,\mathrm{d}s \leq 32\|\kappa\|_\infty|\mathcal{R}|T\|x\|e^{2\|\kappa\|_\infty|\mathcal{R}|T}.$$

and, with Gronwall's Inequality, we obtain

$$\sup_{x\in\mathcal{S},x\neq 0}\mathbb{E}\left(X_x^*(T)^2\right) \leq \sup_{x\in\mathcal{S},x\neq 0}\mathbb{E}\left(Z_x^*(T)^2\right) < +\infty.$$

The family of random variable $(X_x^*(T))$ is uniformly integrable. The proposition is proved. □

PROPOSITION 2.10. *If $(X(t))$ is the Markov process associated to a binary CRN network then the family of random variables*

$$\left(\overline{X}(t)\right) \stackrel{\text{def.}}{=} \left(\frac{1}{\|x\|}X(t/\|x\|)\right),$$

is tight when $x{=}X(0){\in}\mathbb{N}^n\setminus\{0\}$ goes to infinity and any of its limiting point $(\ell(t))$ is a continuous process satisfying the ODE

$$\dot{\ell}(t) = \sum_{\substack{r\in\mathcal{R}\\ y_r^-\in J_2}} \kappa_r \ell_{s_{y_r^-,1}}(t)\ell_{s_{y_r^-,2}}(t)(y_r^+-y_r^-). \tag{2.26}$$

PROOF. This is a simple consequence of Proposition 2.7. □

The timescale $(t/\|x\|)$ and the space scale $1/\|x\|$ are valid for all binary CRNs from the point of view of tightness properties. It does not mean that they are the only ones, or the most meaningful. As it will be seen in Section 5.1, depending on the type of initial state, it may happen that the timescales $(t/\sqrt{\|x\|})$ or (t) and the space scales $1/\sqrt{\|x\|}$ or 1 are appropriate for the analysis of the asymptotic behavior of the time evolution of the CRN. The timescale $(t/\|x\|)$ is well-suited when there are complexes of size two and when the associated chemical species are all in "large" number, of the order of $\|x\|$. Otherwise, it may be too slow to change the state of the CRN, so that a faster timescale has to be used.

5.1. Triangular Binary Networks. We now consider a binary CRN with two chemical species, $m{=}2$, and three distinct complexes C_i, $i{=}1$, 2, 3, of size ≤ 2 and the set of routes is

$$\mathcal{R}{=}\{(C_1,C_2),(C_2,C_3),(C_3,C_1)\}.$$

The purpose of this section is essentially pedagogical, to show, in a simple setting, how the ideas of Sections 3 and 4 can be used in practice, on stability and scaling properties.

As a side result, in Section 2.D of the appendix, a proof of the positive recurrence of a general class of triangle topologies is given: an arbitrary set of chemical species, an arbitrary set of three complexes, i.e. whose sizes are general, and a set of reactions $\mathcal{R}$ containing the three ones from above. Proposition 2.34 of the appendix for triangle topologies can be seen as the analogue of Theorem 7.6.1 of Feinberg [27] for star networks. To the best of our knowledge there are few such stability

results with arbitrary complexes. We have not been able to generalize this proof to a CRN with more than three complexes.

(A) Triangle T1 (B) Triangle T2

FIGURE 2. Triangular CRNs

Note that the CRN T1 of Figure 2 is weakly reversible, with deficiency zero, hence Anderson et al. [7] shows that its associated Markov process is positive recurrent with invariant distribution

$$\pi(x,y) \stackrel{\text{def.}}{=} \frac{1}{Z}\frac{\rho_1^x}{x!}\frac{\rho_2^y}{y!}, \quad (x,y)\in\mathcal{S}, \tag{2.27}$$

where $\mathcal{S}\subset\mathbb{N}^2$ is the state space, $\mathcal{S}=\mathbb{N}^2\backslash\{(0,0),(1,0)\}$, Z is the normalization constant and $\rho_1=\kappa_2/\kappa_{12}$, $\rho_2=\kappa_1/\kappa_{12}$.

The stability analysis of this CRN is revisited by using the criterion of Theorem 2.3 and one of its consequences. This is, of course, not a new result for T1, but this analysis can be done without extra-cost when each chemical species has, in addition, an external source, as in the CRN T2 of Figure 2. This CRN is not anymore weakly reversible, the results of [7] cannot be applied.

The associated Markov process is denoted by $(X^N(t))=(X_1^N(t),X_2^N(t))$. We denote by $x^N=(x_1^N,x_2^N)$ the initial state, which is of norm N, $N=x_1^N+x_2^N$, it is assumed that

$$\lim_{N\to+\infty}\left(\frac{x_1^N}{N},\frac{x_2^N}{N}\right)=(\alpha_1,1-\alpha_1), \tag{2.28}$$

with $\alpha_1\in[0,1]$.

The scalings consider three types of regions of $\mathbb{N}^2$ for the initial state: when the order of magnitude of the two coordinates are respectively of the order of (N,N), $(O(\sqrt{N}),N)$, or $(N,O(1))$. It is shown that starting from a "large" state, three timescales play a role depending on the asymptotic behavior of the initial state:

(a) $t\mapsto t/N$, when both components of the initial state are of the order of N, i.e. when $0<\alpha_1<1$;
(b) $t\mapsto t/\sqrt{N}$, when $\alpha_1=0$ and x_1^N is at most of the order of $\sqrt{N}$.
(c) $t\mapsto t$, when $\alpha_1=1$ and x_2^N is bounded by some constant K.

The boundary effects mentioned in Section 2 play a role in case c), the second coordinate remains in the neighborhood of the origin essentially.

For each of the three regimes, the scaled norm of the state is decreasing to 0, which is helpful to prove positive recurrence. The limit results show additionally that the orders of magnitude in N of both coordinates do not change. In other words the space scale is natural and not the consequence of a specific choice of the timescale. The following proposition gives a formal statement of these assertions.

PROPOSITION 2.11 (Scaling Analysis). *Under the assumptions* (2.28) *on the initial state of the CRN T1 of Figure 2,*

(a) *if* $\alpha_1>0$*, then for the convergence in distribution,*

$$\lim_{N\to+\infty}\left(\frac{X_1^N(t/N)}{N},\frac{X_2^N(t/N)}{N}\right)=(x_{a,1}(t),x_{a,2}(t)), \tag{2.29}$$

where $(x_{a,1}(t),x_{a,2}(t))=(\alpha_1,(1-\alpha_1)\exp(-\kappa_{12}\alpha_1 t))$.

(b) *If* $\alpha_1=0$ *and*

$$\lim_{N\to+\infty}\frac{x_1^N}{\sqrt{N}}=\beta\in\mathbb{R}_+,$$

then, for the convergence in distribution

$$\lim_{N\to+\infty}\left(\frac{X_1^N(t/\sqrt{N})}{\sqrt{N}},\frac{X_2^N(t/\sqrt{N})}{N}\right)=(x_{b,1}(t),x_{b,2}(t)), \tag{2.30}$$

where $(x_{b,1}(t),x_{b,2}(t))$ *is the solution of the ODE*

$$\dot{x}_{b,1}(t)=\kappa_2 x_{b,2}(t),\quad \dot{x}_{b,2}(t)=-\kappa_{12}x_{b,1}(t)x_{b,2}(t), \tag{2.31}$$

with $(x_{b,1}(0),x_{b,2}(0))=(\beta,1)$.

(c) *If the initial state is* $x^N=(N,k)$*, for* $k\in\mathbb{N}$*, then, for the convergence in distribution,*

$$\lim_{N\to+\infty}\left(\frac{X_1^N(t)}{N}\right)=(x_{c,1}(t))\overset{\text{def.}}{=}\left(e^{-\kappa_1 t}\right). \tag{2.32}$$

Proof. See Section 2.C.1 of the appendix. □

Stability Properties. We consider the CRN T2 of Figure 2, with external inputs for both species. Here theorem 2.31 doesn't apply, since the CRN is not weakly reversible. We want to show a stability result on this CRN. It is not difficult to see that the results of Proposition 2.11 holds in this case too since on a finite time interval the number of external arrivals is almost surely finite and independent of N.

Proposition 2.12. *The Markov process associated to the CRN T2 of Figure 2 is positive recurrent.*

Proof. Theorem 2.3 is used. We have to define the stopping time depending on the initial state. Proposition 2.11 does not give a full partitioning of the possible "large" initial states, some additional work has to be done. We ignore the external arrivals, it is easily seen that similar arguments can be used. With high probability, the stopping times chosen are smaller than the instant of the first external arrival. There are three cases.

(1) The initial state is such that

$$\lim_{N\to+\infty}\frac{x_1^N}{\sqrt{N}}\geq 1\quad\text{and}\quad x_2^N\geq 1.$$

We take t_1^N, the instant of the first jump of $(X_N(t))$. Elementary calculations give the relation

$$\limsup_{N\to+\infty}\mathbb{E}_{x^N}(\|X^N(t_1^N)\|)-\|x^N\|\leq-\frac{\kappa_{12}}{\kappa_{12}+\kappa_1},\ \text{and}\ \lim_{N\to+\infty}\sqrt{N}\mathbb{E}\left(t_1^N\right)=0.$$

(2) The initial state is $x^N=(N,0)$.
Let, for $k_0\in\mathbb{N}$, τ_{k_0} be the instant when the (k_0+1)th element S_1 is transformed into an S_2. The norm of the process decreases when some of these molecules of S_2 disappear with reaction $S_1+S_2\rightharpoonup S_1$. The probability for a molecule of S_2 to be destroyed before a new transformation of S_1 into S_2 is lower bounded by $p=\kappa_{12}/(\kappa_1+\kappa_{12})$, therefore in average there are

more than pk_0 molecules of S_2 killed before τ_{k_0}. The reaction $S_2 \rightharpoonup S_1+S_2$ could also create some molecules during the time interval $[0, \tau_{k_0}]$. Its rate being bounded by $\kappa_2 k_0$, and noting that $\mathbb{E}(\tau_{k_0}) \leq k_0/(\kappa_1(N-k_0))$, it is not difficult to show that there exists a constant C_0 such that the relation

$$\mathbb{E}\left(\|X^N(\tau_{k_0})\|\right) \leq N - k_0 p + k_0 \frac{C_0}{N-k_0},$$

holds.

(3) The initial state is $x^N=(x_1^N, N)$ with,

$$\lim_{N\to+\infty} \frac{x_1^N}{\sqrt{N}} \leq 1.$$

We will use the result (b) of Proposition 2.11 and its proof in the appendix (for the convergence of the averages), to obtain the convergence

$$\lim_{N\to+\infty} \frac{1}{N}\mathbb{E}_{x^N}(\|X(1/\sqrt{N})\|) = x_{b,2}(1) < 1.$$

We only have to use Theorem 2.3 to conclude the proof of the proposition. □

6. Agazzi and Mattingly's CRN

In this section, we study the chemical reaction network introduced by Agazzi and Mattingly [4],

$$\emptyset \xrightarrow{\kappa_1} S_1+S_2, \quad S_2 \xrightarrow{\kappa_2} \emptyset, \quad pS_1+qS_2 \xrightarrow{\kappa_3} (q+1)S_2 \xrightarrow{\kappa_4} qS_2, \tag{2.33}$$

for p, $q \in \mathbb{N}$, $p>2$ and $q \geq 2$. In Agazzi and Mattingly [4], the constants considered are $p=5$ and $q=2$.

The purpose of this reference is of showing that with a small modification of the topology of this CRN, its associated Markov process can be positive recurrent, null recurrent, or transient. The main technical part of the paper is devoted essentially to the construction of a Lyapunov function satisfying Condition (2.1).

The continuous time Markov jump process $(X(t))=(X_1(t), X_2(t))$ associated to CRN (2.33) has a Q-matrix given by, for $x \in \mathbb{N}^2$,

$$x \longrightarrow x + \begin{cases} e_1+e_2 & \kappa_1, \\ -pe_1+e_2 & \kappa_3 x_1^{(p)} x_2^{(q)}, \\ -e_2 & \kappa_2 x_2 + \kappa_4 x_2^{(q+1)}, \end{cases}$$

where e_1, e_2 are the unit vectors of $\mathbb{N}^2$. This process is clearly irreducible on $\mathbb{N}^2$, and non explosive since $\emptyset \rightharpoonup S_1+S_2$ is the only reaction increasing the total number of molecules. See the proof of Proposition 2.9 for example. The fact that only this reaction increases the norm of the state suggests that the proof of the positive recurrence should not be an issue.

To prove this positive recurrence, see the proposition below, Agazzi and Mattingly [4] use several energy functions on $\mathbb{N}^2$, which are polynomial functions in x_1 and x_2. The main technical difficulty is of gluing these functions in order to have a global Lyapunov function for which the classical Forster-Lyapunov theorem can be used. Note that there are also interesting null recurrence and transience properties in this reference.

PROPOSITION 2.13. *If $p>2$ and $q \geq 2$, then the Markov process associated to the CRN (2.33) is positive recurrent.*

PROOF. Theorem 2.3 is used with a simple energy function, the norm $\|x\|=x_1+x_2$ of the state $x=(x_1, x_2) \in \mathbb{N}^2$. If the norm of the initial state is large enough, then

the expected value of the norm of the process taken at a convenient stopping time will be smaller, so that Condition (2.16) of Theorem 2.3 holds.

Step 1. As before, for $n\geq 1$, t_n denotes the instant of the nth jump.

$$\mathbb{E}_x\left(\|X(t_1)\|-\|x\|\right)=\left(2\kappa_1-\kappa_2x_2-(p-1)\kappa_3x_1^{(p)}x_2^{(q)}-\kappa_4x_2^{(q+1)}\right)\mathbb{E}_x\left[t_1\right],$$

and, clearly, $\mathbb{E}_x(t_1)\leq 1/\kappa_1$.

If either $x_2\geq K_1=1+2\kappa_1/\kappa_2$ or $q\leq x_2<K_1$ and $x_1\geq K_2=1+2\kappa_1/((p-1)\kappa_3 q!)$, then

$$\mathbb{E}_x\left(\|X(t_1)\|-\|x\|\right)\leq -\gamma\mathbb{E}_x(t_1),$$

for some $\gamma>0$. Condition (2.16) holds for this set of initial states.

Step 2. Now we consider initial states of the form $x_0^N=(N,b)$ with $b<q$ and N large. The third and fourth reactions cannot occur until the instant

$$\tau_1\stackrel{\text{def.}}{=}\inf\{t>0:X_2(t)\geq q\}.$$

Until time τ_1, the process $(X_2(t))$ has the sample path $(L(t))$ of an $M/M/\infty$ queue, see Section 2.3, with arrival rate κ_1 and service rate κ_2. At time τ_1 the state of the process has the same distribution as the random variable

$$(N+\mathcal{N}_{\kappa_0}(0,\tau_1),q),$$

where $\mathcal{N}_{\kappa_0}$ is a Poisson process with rate κ_0. Clearly τ_1 is integrable as well as the random variable $\mathcal{N}_{\kappa_0}(0,\tau_1)$. We have also $X_1(\tau_1\wedge t)=N+\mathcal{N}^1_{\kappa_1}([0,t\wedge\tau_1])$, so $\mathbb{E}_{(N,b)}[X_1(\tau_1)]\leq N+\kappa_1C_1$, for some constant C_1.

To summarize, starting from the initial state $x_0^N=(N,b)$ with $b<q$, the quantities $\mathbb{E}_{x_0^N}(\tau_1)$ and $\mathbb{E}_{x_0^N}(X_1(\tau_1))-N$ are bounded by a constant. We are thus left to study the following case.

Step 3. The initial state is $x_0^N=(N,q)$ with N large.

As long as $X_2(t)\geq q$, the third reaction is active, p copies of S_1 are removed and a copy of S_2 is created. Initially its rate is of the order of N^p, the fastest reaction rate by far. We define ν as the number of jumps before another reaction takes place.

$$\nu\stackrel{\text{def.}}{=}\inf\{n\geq 1:X(t_n)-X(t_{n-1})\neq(-p,1)\},$$

$$\mathbb{P}(\nu>k)=\prod_{i=0}^{k-1}\left(1-\frac{\kappa_1+\kappa_2(q+i)+\kappa_4(q+i)^{(q+1)}}{\kappa_3(N-pi)^{(p)}(q+i)^{(q)}+\kappa_1+\kappa_2(q+i)+\kappa_4(q+i)^{(q+1)}}\right),$$

with the convention that $q^{(q+1)}=0$. For $i\geq 1$,

$$\begin{aligned}&\frac{\kappa_1+\kappa_2(q+i)+\kappa_4(q+i)^{(q+1)}}{\kappa_3(N-pi)^{(p)}(q+i)^{(q)}+\kappa_1+\kappa_2(q+i)+\kappa_4(q+i)^{(q+1)}}\\&\qquad\leq\frac{(\kappa_1+\kappa_2(q+i))(q+i)^{-(q+1)}+\kappa_4}{\kappa_3(N-pi)^{(p)}/i+(\kappa_1+\kappa_2(q+i))(q+i)^{-(q+1)}+\kappa_4}\leq\frac{iC_0}{(N-pi)^{(p)}+iC_0},\end{aligned}$$

for some apropriate constant $C_0>0$. Hence, if we fix $0<\delta<1/2p$,

$$\mathbb{E}_{x_0^N}(\nu)\geq\delta N\mathbb{P}(\nu>\delta N)\geq\delta N\left(1-\frac{\delta NC_0}{(N-p\lfloor\delta N\rfloor)^{(p)}+\delta NC_0}\right)^{\lfloor\delta N\rfloor},$$

so that, since $p>2$,

$$\liminf_{N\to+\infty}\frac{1}{N}\mathbb{E}_{x_0^N}(\nu)\geq\delta. \tag{2.34}$$

We define $\tau_2{=}t_\nu$, obviously

$$\mathbb{E}_{x_0^N}(\tau_2) \leq \frac{1}{\kappa_1},$$

and we have

$$\mathbb{E}_{x_0^N}(\|X(\tau_2)\|{-}\|x_0^N\|) \leq (1{-}p)\mathbb{E}_{x_0^N}(\nu){+}2 \leq -\gamma N,$$

for some $\gamma{>}0$ if N is sufficiently large, using Relation (2.34). Consequently is easy to see that there is a convenient constant K such that Condition (2.16) holds for this set of initial states and the stopping time τ_2, and also for the initial states of Step 2 and the stopping time $\tau_1{+}\tau_2{\circ}\theta_{\tau_1}$. The proposition is proved. □

A Scaling Picture. The key argument of the proof of the positive recurrence is somewhat hidden behind an estimate of the expected value of the hitting time ν in Step 3. It is not difficult to figure out that, starting from the state (N, q), the "right" timescale is $t{\mapsto}t/N^{p+q-1}$. In this section we sketch a scaling argument to describe in more detail how the norm of the state goes to 0. It could also give an alternative way to handle Step 3.

Define the Markov jump process $(Z_N(t)) = (Z_1^N(t), Z_2^N(t))$ corresponding to the last two reactions of the CRN network (2.33). Its Q-matrix is given by, for $z{\in}\mathbb{N}^2$,

$$z \longrightarrow z{+}\begin{cases} -pe_1{+}e_2 & \kappa_3 z_1^{(p)} z_2^{(q)}, \\ -e_2 & \kappa_4 z_2^{(q+1)}, \end{cases} \tag{2.35}$$

with initial state (N, q). The scaling results of this section are obtained for this process. It is not difficult to show that they also hold for the CRN network (2.33) since the discarded reactions are on a much slower timescale.

Define the Markov jump process $(Y_N(t)) = (Y_1^N(t), Y_2^N(t))$ whose Q-matrix is given by, for $y{\in}\mathbb{N}^2$,

$$y \longrightarrow y{+}\begin{cases} -pe_1{+}e_2 & \kappa_3 y_1^{(p)}, \\ -e_2 & \kappa_4(y_2{-}q), \end{cases}$$

with the same initial state. If $p{\geq}2$, with the same arguments as in the proof of Proposition 2.11 (see Section 2.C.1 of the appendix), it is not difficult to show the convergence in distribution

$$\begin{aligned} \lim_{N\to+\infty} \left(\frac{1}{N}\left(Y_1^N, Y_2^N\right)\left(\frac{t}{N^{p-1}}\right)\right) &= (y_1(t), y_2(t)) \\ &\stackrel{\text{def.}}{=} \left(\frac{1}{\sqrt[p-1]{p(p{-}1)\kappa_3 t{+}1}}, \frac{1{-}y_1(t)}{p}\right) \end{aligned} \tag{2.36}$$

From this convergence we obtain that for any $\eta{\in}(0, 1/p)$, the hitting time $H_Y^N(\eta)$ of $\lfloor \eta N \rfloor$ by $(Y_2^N(t))$ is such that $(N^{p-1}H_Y^N(\eta))$ converges in distribution to some constant.

For $t{\geq}0$, define the stopping time

$$\tau_t^N = \inf\left\{s{>}0 : \int_0^s \frac{1}{Y_2^N(u)^{(q)}}\,\mathrm{d}u \geq t\right\},$$

and $(\widetilde{Z}^N(t)){=}(Y^N(\tau_t^N))$, then it is easy to check that $(\widetilde{Z}^N(t))$ is a Markov process whose Q-matrix is given by Relation (2.35). See Section III.21 of Rogers and Williams [68] for example. Consequently, $(\widetilde{Z}^N(t))$ has the same distribution as $(Z^N(t))$.

PROPOSITION 2.14. *If p, $q\geq 2$, $(X^N(0))=(\lfloor \delta N\rfloor, \lfloor (1-\delta)N/p\rfloor)$, for some $\delta\in(0,1)$, then for the convergence in distribution*

$$\lim_{N\to+\infty}\left(\frac{1}{N}X^N\left(\frac{t}{N^{p+q-1}}\right)\right)=(x_1(t),x_2(t))=\left(\left(y_1,\frac{1-y_1}{p}\right)\left(\phi^{-1}(t)\right)\right),$$

with

$$(y_1(t))=\left(\frac{\delta}{\sqrt[p-1]{p(p-1)\delta^{p-1}\kappa_3 t+1}}\right) \text{ and } \phi(t)\stackrel{\text{def.}}{=}\int_0^t\frac{p^q}{(1-y_1(s))^q}\,\mathrm{d}s.$$

PROOF. As mentioned above, from this initial state and this timescale, the processes $(Z^N(t))$ and $(X^N(t))$ have the same asymptotic behavior for values of the order of N. The proof uses the convergence (2.36) and the time-change argument described above. □

The above proposition shows that on a convenient timescale, both coordinates of $(X^N(t))$ are of the order of N. The scaled version of the first one is converging to 0, while the second component is increasing.

If $Y^N(0)=(\lfloor \delta N\rfloor, \lfloor (1-\delta)N/p\rfloor)$, for some $\delta>0$, let

$$H^N=\inf\{t>0: Y_1^N(t)\leq \sqrt[p]{N}\}.$$

By writing the evolution of $(Y^N(t))$ in terms of an SDE like Relation (2.4), one easily obtains,

$$\mathbb{E}(Y_1^N(H^N\wedge t))=\lfloor \delta N\rfloor - p\kappa_3\, E\left(\int_0^{H^N\wedge t} Y_1^N(s)^{(p)}\,\mathrm{d}s\right)$$
$$\leq \lfloor \delta N\rfloor - p\kappa_3\left(\sqrt[p]{N}\right)^{(p)} E\left(H^N\wedge t\right),$$

hence, by using the monotone convergence theorem, we obtain that

$$E\left(H^N\right)\leq\frac{\lfloor \delta N\rfloor}{p\kappa_3(\sqrt[p]{N})^{(p)}},\ \text{hence, } \sup_N E\left(H_K^N\right)<+\infty,$$

since $p\geq 2$. It is easily seen that the same property holds for $(X_1^N(t))$.

To finish the description of the return path to $(0,0)$, we can assume therefore that $X^N(0)=(\lfloor \sqrt[p]{N}\rfloor, N)$. It is not difficult to see that the reaction $(q+1)S_2\xrightarrow{\kappa_4} qS_2$ is driving the evolution as long as $(X_2^N(t))$ is "large" since $(X_1^N(t))$ cannot grow significantly on the corresponding timescale. More formally, also with the same arguments as in Section 2.11, the convergence in distribution

$$\lim_{N\to+\infty}\left(\frac{1}{N}\left(X_1^N,X_2^N\right)\left(\frac{t}{N^q}\right)\right)=\left(0,\frac{1}{\sqrt[q]{1+\kappa_4 q t}}\right)$$

holds.

7. A CRN with Slow and Fast Timescales

In this section, the positive recurrence and scaling properties of the following interesting CRN are investigated

$$\emptyset\underset{\kappa_1}{\overset{\kappa_0}{\rightleftharpoons}}S_1+S_2,\qquad pS_1+S_2\underset{\kappa_3}{\overset{\kappa_2}{\rightleftharpoons}}pS_1+2S_2, \tag{2.37}$$

with $p\geq 2$. This is an important reference example introduced and discussed from the point of view of its stability properties in Agazzi et al. [2] for $p=2$. It shows how boundary effects can complicate the verification of a Lyapunov criterion. In [2], the energy function used for the entropy function. We show how Condition 2.2 can be used for positive recurrence and that a scaling analysis gives an interesting insight for the time evolution of this CRN.

This CRN exhibits several distinctive features of chemical reaction networks. It provides an important example of a non-trivial CRN for which the results and ideas of Section 4 and Section 3 can be used together with time change arguments, and all of that within a (quite) limited technical framework. For this reason, we have chosen to develop completely the technical arguments used. The results obtained are interesting in their own right in fact.

Section 7.1 investigates the positive recurrence properties. It is an occasion to have an other look at the choice of a Lyapunov function in view of Condition 2.16 of Theorem 2.3. Section 7.2 considers the limiting behavior of the sample paths of the CRN with a large initial state close to one of the axes. As it can be expected, in both sections boundary effects play a very important role: the second reaction cannot occur if there are less than p copies of S_1, and if the number of copies of S_2 is zero, only external arrivals change the state of the CRN.

The Markov process $(X(t)) = (X_1(t), X_2(t))$ associated to this CRN has a Q-matrix Q given by, for $x \in \mathbb{N}^2$,

$$x \longrightarrow x + \begin{cases} e_1+e_2 & \kappa_0, \\ -e_1-e_2 & \kappa_1 x_1 x_2, \end{cases} \quad x \longrightarrow x + \begin{cases} e_2 & \kappa_2 x_1^{(p)} x_2, \\ -e_2 & \kappa_3 x_1^{(p)} x_2^{(2)}, \end{cases}$$

where e_1, e_2 are the unit vectors of $\mathbb{N}^2$.

By using the SDE formulation of Section 2.4 of the appendix, the associated Markov process can be represented by the solution $(X(t))=(X_1(t), X_2(t))$ of the SDE

$$\begin{cases} \mathrm{d}X_1(t) &= \mathcal{P}_{X,0}((0,\kappa_0),\mathrm{d}t) - \mathcal{P}_{X,1}((0,\kappa_1 X_1 X_2(t-)),\mathrm{d}t), \\ \mathrm{d}X_2(t) &= \mathcal{P}_{X,0}((0,\kappa_0),\mathrm{d}t) - \mathcal{P}_{X,1}((0,\kappa_1 X_1 X_2(t-)),\mathrm{d}t) \\ &\qquad + \mathcal{P}_{X,2}\left(\left(0,\kappa_2 X_1^{(p)} X_2(t-)\right),\mathrm{d}t\right) \\ &\qquad - \mathcal{P}_{X,3}\left(\left(0,\kappa_3 X_1^{(p)} X_2^{(2)}(t-)\right),\mathrm{d}t\right), \end{cases} \tag{2.38}$$

where $\mathcal{P}_{X,i}$, $i \in \{0,1,2,3\}$, are fixed independent Poisson processes on $\mathbb{R}_+^2$ with intensity measure $\mathrm{d}s \otimes \mathrm{d}t$. A notation of this kind $\mathcal{P}_A$ or $\mathcal{P}_{A,i}$ will be used for several A in the following, with the same assumptions on the distribution and the independence properties.

A SLOW RETURN TO 0. The reactions of the second linkage class of this CRN need p copies of S_1 to be active. If the initial state is $(0, N)$, copies of S_1 are created at rate κ_0, but they are removed quickly at a rate greater than $\kappa_1 N$. The first instant when p copies of S_1 are present has an average of the order of N^{p-1}. See Lemma 2.16. At that instant, the number of S_2 species is $N+p$, and the second coordinate can then decrease, quickly in fact. The network exhibits a kind of bi-modal behavior due to this boundary condition.

Starting from the initial state $x=(0,N)$, the time to decrease $(X_2(t))$ by an amount of the order of N has thus an average of the order of $\|N\|^{p-1}$. When $p>2$ and if we take the usual norm $\|\cdot\|$ as a Lyapunov function, this results is at odds with one of the conditions of positive recurrence criterion of Proposition 2.6. This problem could in fact be fixed at the cost of some annoying technicalities. Our approach will be of taking another simple, and somewhat natural, Lyapunov function. See Section 7.1. An initial state of the form $(N, 0)$ leads also to another interesting boundary behavior.

7.1. Positive Recurrence.

PROPOSITION 2.15. *The Markov process $(X(t))$ is positive recurrent.*

Theorem 2.3 is used to prove this property. The proof is not difficult but it has to be handled with some care. We will introduce two auxiliary processes with

which the process $(X_N(t))$ can be decomposed. One describes the process when the first coordinate is below p and the other when the second coordinate is greater or equal to 1. This representation gives a more formal description of the bi-modal behavior mentioned above. Additionally, it will turn out to be helpful to establish the scaling properties of this CRN in Section 7.2. For $x{=}(x_1,x_2){\in}\mathbb{N}^2$, we introduce

$$f_p(x){=}x_1{+}x_2^p, \tag{2.39}$$

f_p will be our Lyapunov function. The strategy is of analyzing separately the two boundary behaviors. The first one is essentially associated with the initial state $(0,N)$ which we have already seen. The other case is for an initial state of the form $(N,0)$, the problem here is of having the second coordinate positive sufficiently often so that reaction $S_1{+}S_2 \rightharpoonup \emptyset$ can decrease significantly the first coordinate.

7.1.1. *Large Initial State along the Horizontal Axis.* In this section it is assumed that the initial state is $x(0){=}(x_1^0,b)$, where $b{\in}\mathbb{N}$ is fixed and x_1^0 is "large". Without loss of generality one can assume $b{>}0$, otherwise nothing happens until an external arrival.

As long as the second coordinate of $(X(t))$ is non-null the transitions associated to $\mathcal{P}_{X,i}$, $i{=}2$, 3 occur at a fast rate. When $(X_2(t))$ is 0, only one chemical reaction may happen, external arrivals and at a "slow" rate κ_0.

We define by induction the non-increasing sequence (T_k) as follows, $T_0{=}0$, and

$$T_{k+1} = \inf\{t{>}T_k : X_1(t){-}X_1(t{-}){=}{-}1\}.$$

The variables (T_k) are stopping times for the underlying filtration $(\mathcal{F}_t)$ defined as in the appendix, see Relation (2.11).

For $t{>}0$, by using the fact that the Poisson process $\mathcal{P}_{X,i}$, $i{=}1$, 2, 3 are independent and $(X_2(t))$ is greater than 1 until T_1 at least, we have

$$\mathbb{P}(T_1{\geq}t) \leq \mathbb{E}\left(\exp\left(-\kappa_1 x_1^0 \int_0^t X_2(s)\,\mathrm{d}s\right)\right) \leq \exp\left(-\kappa_1 x_1^0 t\right),$$

hence $\mathbb{E}(T_1){\leq}1/(\kappa_1 x_1^0)$. Similarly, with the strong Markov property, for $1{\leq}k{<}x_1^0$,

$$\mathbb{E}(T_{k+1}{-}T_k){\leq}\frac{1}{\kappa_0}{+}\frac{1}{\kappa_1(x_1^0{-}k)},$$

the additional term $1/\kappa_0$ comes from the fact that $X_2(T_k)$ can be zero, so that one has to wait for an exponential time with parameter κ_0 to restart the CRN.

For $n_0{\geq}1$, we have seen that the random variable T_{n_0} is stochastically bounded by the sum of $2n_0$ i.i.d. exponentially distributed random variables with some positive rate, hence

$$C_0 \stackrel{\text{def.}}{=} \sup_{x_1^0 > n_0} \mathbb{E}_{x_1^0}(T_{n_0}) < +\infty$$

Let $\mathcal{E}_1$ be the event when $\mathcal{P}_{X,1}$ has a jump before $\mathcal{P}_{X,0}$ in SDE (2.38), then

$$\mathbb{P}(\mathcal{E}_1^c){\leq}\frac{\kappa_0}{\kappa_1 x_1^0{+}\kappa_0}.$$

Similarly, for $k{\geq}2$, $\mathcal{E}_k$ is a subset of the event $\mathcal{E}_{k-1}$ for which $\mathcal{P}_{X,1}$ has a jump before $\mathcal{P}_{X,0}$ after the first time after T_k when $(X_2(t))$ is greater than 1, then

$$\mathbb{P}_{x_1^0}(\mathcal{E}_k^c){\leq}\sum_{i=0}^{k-1}\frac{\kappa_0}{\kappa_1(x_1^0{-}i){+}\kappa_0}{\leq}\frac{\kappa_0 k}{\kappa_1(x_1^0{-}k){+}\kappa_0}. \tag{2.40}$$

Let s_1 be the first instant of jump of $\mathcal{P}_{X,0}((0,\kappa_0){\times}(0,t])$. From $t{=}0$, as long as the point process $\mathcal{P}_{X,0}$, does not jump in SDE (2.38), that is, on the time interval

$[0, s_1]$, up to a change of time scale $t \mapsto X_1 X_2(t)$, the process $(X_1(t), X_2(t))$ has the same sequence of visited states as the solution $(Y(t))$ of the SDE

$$\begin{cases} \mathrm{d}Y_1(t) &= -\mathcal{P}_{X,1}((0,\kappa_1),\mathrm{d}t), \\ \mathrm{d}Y_2(t) &= -\mathcal{P}_{X,1}((0,\kappa_1)),\mathrm{d}t) \\ &\quad +\mathcal{P}_{Y,2}\left(\left(0,\kappa_2 Y_1(t-)^{(p)-1}\right),\mathrm{d}t\right) \\ &\quad -\mathcal{P}_{Y,3}\left(\left(0,\kappa_3 Y_1^{(p)-1}(Y_2(t-)-1)^+\right),\mathrm{d}t\right), \end{cases} \tag{2.41}$$

with the same initial state and the slight abuse of notation $y^{(p)-1}=y^{(p)}/y$. In particular if u_1 is the first instant when $(Y_1(t))$ has a downward jump, an independent exponential random variable with parameter κ_1, then the relation $Y_2(u_1)=X_2(T_1)$ holds on the event $\{T_1 \leq s_1\}$.

From $t=0$, as long as the first coordinate of $(Y_1(t))$ does not change, the second component $(Y_2(t))$ has the same distribution as $(L_b((x_1^0)^{(p)-1}t))$, where $(L_b(t))$ is a birth and death process with birth rate κ_2 and death rate $\kappa_3(x-1)$ in state $x \geq 1$ and initial state b. It is easily seen that it is a positive recurrent Markov process and that $(\mathbb{E}(L_b(t)^p))$ is a bounded function. Consequently,

$$\sup_{x(0)} E\left(X_2(T_1)^p\right) \leq C_1 < +\infty, \tag{2.42}$$

by induction, the same result holds for T_{n_0} for a convenient constant C_1.

Note that if $X_2(T_1-)=1$, the next reaction happening after T_1 will be $\emptyset \rightharpoonup S_1 + S_2$, and therefore the jump downward of X_1 will be canceled. A decrease at time T_1 of $(X_1^N(t))$ is sustainable if it happens when $X_2(T_1-) \geq 2$ i.e. if $L_b((x_1)^{(p)-1}u_1) \neq 0$. It is not difficult to construct a coupling with $(L_0(t))$, a birth and death process starting at 0, such that $L_b(t) \geq L_0(t)$ holds for all $t \geq 0$. The convergence of $(L_0(t))$ to equilibrium gives the existence of $K_0 \geq 0$ and $\eta_0 > 0$ such that if $x_1^0 \geq K_0$ then $\mathbb{P}(L_0((x_1)^{(p)-1}u_1) > 0) \geq \eta_0$.

We can gather these results, and the stochastic bound on T_{n_0}, to get the relations

$$\begin{aligned} E_{x(0)}(f_p(X(T_{n_0}))) - f_p(x(0)) &\leq -n_0\eta_0 \mathbb{P}_{x(0)}(\mathcal{E}_{n_0}) \\ &\quad + \mathbb{E}_{x(0)}\left(\mathcal{P}_{X,0}\left((0,\kappa_0)\times(0,T_{n_0}]\mathbb{1}_{\{\mathcal{E}_{n_0}^c\}}\right)\right) + E\left(X_2(T_{n_0})^p\right) - b^p \\ &\leq -\eta_0 n_0 + n_0\eta_0 \mathbb{P}_{x(0)}(\mathcal{E}_{n_0}^c) + \kappa_0 C_0 + C_1. \end{aligned}$$

One first choose n_0 so that $n_0 > 3(\kappa_0 C_0 + C_1)/\eta_0$ and then with Relation (2.40), $K_1 \geq K_0$ such that $n_0\eta_0 \mathbb{P}_{K_1}(\mathcal{E}_k^c) < (\kappa_0 C_0 + C_1)$. We obtain therefore that if $x_1^0 > K_1$, then

$$\mathbb{E}_{x(0)}\left(f_p\left(X(T_{n_0})\right) - f_p(x(0))\right) \leq -\delta, \tag{2.43}$$

for some $\delta > 0$ and $\sup(\mathbb{E}_{x(0)}(T_{n_0}) : x_1 \geq K) < +\infty$. Relation (2.43) shows that Condition (2.16) of Theorem (2.3) is satisfied for our Lyapunov function f_p and stopping time T_{n_0} for the initial state of the form (x_1^0, b).

7.1.2. *Initial State with a Large Second Component.* In this section it is assumed that the initial state is $x(0)=(a, x_2^0)$ with $a < p$ and x_2^0 is large. We note that, as long as $(X_1(t))$ is strictly below p, the two coordinates experience the same jumps, the quantity $(X_2(t) - X_1(t))$ does not change. For this reason, for $x \geq 0$ and $k \leq p-1$, we introduce a process $(Z(k,x,t))$ which will be used to express $(X(t))$ when its first coordinate is less than $p-1$. It is the solution of the SDE

$$\mathrm{d}Z(k,x_2^0,t) = \mathcal{N}_{\kappa_0}(\mathrm{d}t) - \mathcal{P}_Z((0,\kappa_1 Z(k,x_2^0,t-)(x_2^0 - k + Z(k,x_2^0,t-)))),\mathrm{d}t), \tag{2.44}$$

with $Z(k,x_2^0,0)=k$ and $\mathcal{P}_Z$ is a Poisson process on $\mathbb{R}_+^2$.

Setting for $z < p$

$$T_Z(z, x_2^0) \stackrel{\text{def.}}{=} \inf\{t{>}0 : Z(z, x_2^0, t){=}p\},$$

if $X(0){=}(0, x_2^0)$, then it is easily seen that the relation

$$(X(t{\wedge}T_Z(0, x_2^0))) \stackrel{\text{dist.}}{=} (Z(0, x_2^0, t{\wedge}T_Z(0, x_2^0)), x_2^0{+}Z(0, x_2^0, t{\wedge}T_Z(0, x_2^0)))$$

holds by checking the jump rates.

We define, for $x{=}(x_1, x_2){\in}\mathbb{N}^2$,

$$\lambda(x) = \kappa_0{+}\kappa_1 x_1 x_2{+}\kappa_2 x_1^{(p)} x_2{+}\kappa_3 x_1^{(p)} x_2^{(2)},$$

it is the total jump rate of $(X(t))$ in state x.

LEMMA 2.16. *For $x_1^0{\geq}\kappa_0/(\kappa_1 p)$,*

$$\limsup_{x_2^0 \to +\infty} \frac{\mathbb{E}(T_Z(0, x_2^0))}{(x_2^0)^{p-1}} \leq C_2,$$

for some constant C_2.

PROOF. A simple coupling shows that the process $(Z(0, x, t)$ stopped at time $T_Z(0, x)$ is lower bounded by a birth and death process $(U(t))$ starting at 0 with, in state x, a birth rate κ_0 and a death rate $a_1{=}\kappa_1 p(x + p)$. Denote by H the hitting time of p by $(U(t))$, then it is easily seen, that, for $0{<}k{<}p$,

$$(\mathbb{E}_k(H){-}\mathbb{E}_{k+1}(H)) = \frac{a_1}{\kappa_0}(\mathbb{E}_{k-1}(H){-}\mathbb{E}_k(H)) + \frac{1}{\kappa_0},$$

with $\mathbb{E}_0(H){-}\mathbb{E}_1(H){=}1/\kappa_0$.In particular $\mathbb{E}(T_Z(0, x)){\leq}\mathbb{E}_0(H)$. We derive the desired inequality directly from this relation. □

(a) If $x_1{\geq}p$.
Define

$$C_1 \stackrel{\text{def.}}{=} \sup_{x_2 \geq 1} \left(\frac{(x_2{+}p)^{(p)}{-}(x_2)^{(p)}}{x_2^{p-1}} \right) < +\infty$$

and

$$\tau_1 \stackrel{\text{def.}}{=} \inf\{t{>}0 : \Delta X_1(t){+}\Delta X_2(t) \neq -1\},$$

where $\Delta X_i(t){=}X_i(t){-}X_i(t{-})$, for $i{\in}\{1, 2\}$ and $t{\geq}0$. The variable τ_1 is the first instant when a reaction other than $pS_1{+}2S_2 \rightharpoonup pS_1{+}S_2$ occurs.

For $1{\leq}k_0{<}x_2$, then

$$\mathbb{P}_{x(0)}(X_2(\tau_1) \leq x_2{-}k_0 - 1) \geq \prod_{i=0}^{k_0} \frac{\kappa_3 x_1^{(p)}(x_2{-}i)^{(2)}}{\lambda((x_1, x_2{-}i))} \geq p_{k_0} \stackrel{\text{def.}}{=} \prod_{i=0}^{k_0} \frac{\kappa_3 p^{(p)}(x_2{-}i)^{(2)}}{\lambda((p, x_2{-}i))}$$

and there exists $K_{k_0}{\geq}k_0$ such that if $x_2{\geq}K_{k_0}$, then

$$(x_2{-}k_0)^{(p)}{-}(x_2)^{(p)} \leq -\frac{k_0}{2} x_2^{p-1} \text{ and } p_{k_0} \geq \frac{1}{2},$$

from these relations, we obtain the inequality

$$\begin{aligned} \mathbb{E}_{x(0)}\left(f_p\left(X(\tau_1)\right){-}f_p(x)\right) &\leq 1{+}\left((x_2{-}k_0)^{(p)}{-}x_2^{(p)}\right) p_{k_0}{+}\left((x_2{+}1)^{(p)}{-}x_2^{(p)}\right) \\ &\leq \left(-\frac{k_0}{4}{+}1{+}C_1\right) x_2^{p-1}. \end{aligned} \tag{2.45}$$

We choose $k_0{=}\lceil 4(3{+}2C_1)\rceil$, hence, for $x_2{\geq}K_{k_0}$ the relation

$$\mathbb{E}_{x(0)}\left(f_p\left(X(\tau_1)\right){-}f_p(x)\right) \leq -2x_2^{p-1}.$$

holds, and note that $\mathbb{E}(\tau_1){\leq}1/\kappa_0$.

(b) If $x_1 \leq p-1$.
Define

$$\tau_0 = \inf\{t>0 : X_1(t) \geq p\},$$

When $x_1 = 0$, the variable τ_0 has the same distribution as $T_Z(0, x_2)$, otherwise it is easily seen that $\mathbb{E}_{x(0)}[\tau_0] \leq \mathbb{E}[T_Z(0, x_2)]$. Lemma 2.16 gives therefore a constant $C_2 > 0$ so that

$$\sup_{x_2 \geq K_{k_0}} \left(\frac{\mathbb{E}_x(\tau_0)}{x_2^{p-1}} \right) < C_2.$$

The state of the process at time τ_0 is $X(\tau_0)=(p, x_2+(p-x_1))$, in particular

$$\mathbb{E}_{x(0)} \left(f_p(X(\tau_0)) - f_p(x) \right) \leq p + C_1 x_2^{p-1},$$

and at that instant, we are in case a).

The convenient stopping time is defined as $\tau_2 \overset{\text{def.}}{=} \tau_0 + \tau_1(\theta_{\tau_0})$. With k_0 and K_{k_0} as before, if $x_2 \geq K_{k_0}$, by using Relation (2.45), we obtain the inequality

$$\begin{aligned}\mathbb{E}_x \left[f(X(\tau_2)) - f(x) \right] &\leq p + C_1 x_2^{p-1} + \mathbb{E}_{(p, x_2+(p-x_1))} \left[f(X(\tau_1)) - f(x) \right] \\ &\leq p + C_1 x_2^{p-1} + \left(-\frac{k_0}{4} + 1 + C_1 \right) (x_2 + (p - x_1))^{p-1} \leq -x_2^{p-1}\end{aligned}$$

holds.

Proof of Proposition 2.15. Theorem 2.3 can be used as a consequence of a), b), and Relation (2.43). □

7.2. A Scaling Picture. We investigate the scaling properties of $(X_N(t))$ when the initial state is of the form $(N, 0)$ or $(0, N)$ essentially. In the first case, an averaging principle is proved on a convenient timescale. A time change argument is an important ingredient to derive the main limiting result. In the second case, the time evolution of the second coordinate of the process is non-trivial only on "small" time intervals but with a "large" number of jumps, of the order of N. This accumulation of jumps has the consequence that the convergence of the scaled process cannot hold with the classical Skorohod topology on $\mathcal{D}(\mathbb{R}_+)$. There are better suited topologies to handle this kind of situation. To keep the presentation simple, we have chosen to work with the weaker topology in the space of random measures, for the occupation measures of the sequence of scaled processes.

7.2.1. *Horizontal Axis.* For $N \geq 1$, the initial state is (x_1^N, b), $b \in \mathbb{N}$ is fixed, it is assumed that

$$\lim_{N \to +\infty} \frac{x_1^N}{N} = \alpha_1 > 0. \tag{2.46}$$

When the process $(X_2(t))$ hits 0, it happens only for a jump of $\mathcal{P}_{X,1}$, all reactions but one are inactive. One has to wait for a jump of $\mathcal{N}_{\kappa_0}$ to restart the activity of the CRN.

We introduce the process $(Y_N(t))=(Y_1^N(t), Y_2^N(t))$, solution of the SDE,

$$\begin{cases} \mathrm{d}Y_1^N(t) &= \mathcal{N}_{\kappa_0}(\mathrm{d}t) - \mathbb{1}_{\{Y_2^N(t-)>1\}} \mathcal{P}_{Y,1}((0, \kappa_1 Y_1^N Y_2^N(t-)), \mathrm{d}t), \\ \mathrm{d}Y_2^N(t) &= \mathcal{N}_{\kappa_0}(\mathrm{d}t) - \mathbb{1}_{\{Y_2^N(t-)>1\}} \mathcal{P}_{Y,1}((0, \kappa_1 Y_1^N Y_2^N(t-)), \mathrm{d}t) \\ &\qquad + \mathcal{P}_{Y,2}\left(\left(0, \kappa_2 (Y_1^N)^{(p)} Y_2^N(t-)\right), \mathrm{d}t \right) \\ &\qquad - \mathcal{P}_{Y,3}\left(\left(0, \kappa_3 (Y_1^N)^{(p)} Y_2^N(t-)^{(2)}\right), \mathrm{d}t \right), \end{cases} \tag{2.47}$$

with initial condition $(Y_1^N(0), Y_2^N(0))=(x_1^N, b)$.

The process $(Y_N(t))$ behaves as $(X(t))$ except that its second coordinate cannot be 0 because the associated transition is excluded. In state $(x,1)$ for $(X(t))$, if the Poisson process $\mathcal{P}_{X,1}$ "rings" in Relation (2.38), the state becomes $(x-1,0)$. It stays in this state for a duration which is exponentially distributed with parameter κ_0 after which the state of $(X(t))$ is back to $(x,1)$. These time intervals during which $(X_2(t))$ is 0 are, in some sense, "wiped out" to give $(Y_N(t))$. This can be expressed rigorously via a time change argument. See Chapter 6 of Ethier and Kurtz [24] for example.

Now the strategy to obtain a scaling result for $(X_1^N(t)$ is of establishing a limit result for $(Y_N(t))$ and, with an appropriate change of timescale, express the process $(X_1^N(t))$ as a "nice" functional of $(Y_N(t))$.

Define

$$\left(\overline{Y}_1^N(t)\right) = \left(\frac{Y_1^N(t)}{N}\right) \text{ and } \langle \mu_N, f\rangle \stackrel{\text{def.}}{=} \int_0^{+\infty} f(s, Y_2^N(s))\,\mathrm{d}s,$$

if f is a function on $\mathbb{R}_+\times\mathbb{N}$ with compact support, μ_N is the *occupation measure* associated to $(Y_2^N(t))$. See Kurtz [55].

PROPOSITION 2.17. *The sequence $(\mu_N, (\overline{Y}_1^N(t)))$ is converging in distribution to a limit $(\mu_\infty, (y_\infty(t))$ defined by*

$$\langle \mu_\infty, f\rangle = \int_{\mathbb{R}_+\times\mathbb{N}} f(s,x)\pi_Y(\mathrm{d}x)\,\mathrm{d}s,$$

if $f\in\mathcal{C}_c(\mathbb{R}_+\times\mathbb{N})$, the function $(y_\infty(t))$ is given by

$$y_\infty(t) = \alpha_1 \exp\left(-\frac{\kappa_1\kappa_2}{\kappa_3}t\right) \text{ for } t\geq 0, \tag{2.48}$$

and π_Y is the distribution on $\mathbb{N}\backslash\{0\}$ defined by, for $x\geq 1$,

$$\pi_Y(x) = \frac{1}{x!}\left(\frac{\kappa_2}{\kappa_3}\right)^x \frac{1}{e^{\kappa_2/\kappa_3}-1}.$$

PROOF. The proof is quite standard. Because of the term $Y_1^N(t)^{(p)}$ in the SDE of the process $(Y_2^N(t))$, the only (small) difficulty is to take care of the fact that $(Y_1^N(t))$ has to be of the order of N, otherwise $(Y_2^N(t))$ may not be a "fast" process. We give a sketch of this part of the proof.

Let a, $b\in\mathbb{R}_+$ such that $0<a<\alpha_1<b$, and

$$S_N = \inf\left\{t>0, \overline{X}_1^N(t)\notin(a,b)\right\}.$$

Let $(L(t))$ a birth and death process on $\mathbb{N}$, when in state $y\geq 1$, its birth rate is βy and the death rate is $\delta y(y-1)$, with $\beta=(\kappa_0+\kappa_2 b^p)$ and $\delta=\kappa_3 a^p$. Its invariant distribution is a Poisson distribution with parameter β/δ conditioned to be greater or equal to 1.

If N is sufficiently large, we can construct a coupling of $(Y_2^N(t))$ and $(L(t))$, with $L(0)=Y_2^N(0)$ and such that the relation

$$Y_2^N(t)\leq L(N^p t)$$

holds for $t\in[0,S_N)$.

For $t>0$,

$$\frac{Y_1^N(t)}{N} \geq \frac{x_1^N}{N} - \kappa_1\int_0^t Y_1^N(s)Y_2^N(s)\,\mathrm{d}s - M_Y^N(t),$$

where $(M_Y^N(t))$ is the martingale given by

$$\left(\frac{1}{N}\int_0^t \mathbb{1}_{\{Y_2^N(s-)>1\}}\left[\mathcal{P}_{Y,1}((0,\kappa_1 Y_1^N(s-)Y_2^N(s-)),\mathrm{d}s) - \kappa_1 Y_1^N(s)Y_2^N(s))\,\mathrm{d}s\right]\right),$$

we have

$$\frac{Y_1^N(t\wedge S_N)}{N} \geq \frac{x_1^N}{N} - \kappa_1 b \int_0^t L(N^p s), \mathrm{d}s + M_Y^N(t\wedge S_N), \tag{2.49}$$

and

$$\left\langle M_Y^N \right\rangle (t\wedge S_N) \leq \frac{b}{N} \int_0^t L(N^p s)\, \mathrm{d}s.$$

By the ergodic theorem applied to $(L(t))$, almost surely

$$\lim_{N\to+\infty} \int_0^t L(N^p s)\, \mathrm{d}s = \lim_{N\to+\infty} \int_0^t \mathbb{E}(L(N^p s))\, \mathrm{d}s$$
$$= \lim_{N\to+\infty} \frac{1}{N^p} \int_0^{N^p t} L(s), \mathrm{d}s = \frac{\beta}{\delta} \frac{t}{1-\exp(-\beta/\delta)}.$$

We deduce that $(M_Y^N(t), t \leq \eta)$ is converging in distribution to 0 by Doob's Inequality and, with Relation (2.49), that there exists $\eta>0$ such that

$$\lim_{N\to+\infty} \mathbb{P}(S_N>\eta) = 1. \tag{2.50}$$

For $\varepsilon>0$ and $K>0$,

$$\mathbb{E}(\mu_N([0,\eta]\times[K,+\infty])) \leq \mathbb{E}\left(\int_0^{\eta\wedge S_N} \mathbb{1}_{\left\{Y_2^N(s)\geq K\right\}}\, \mathrm{d}s \right) + \eta\mathbb{P}(S_N \leq \eta)$$
$$\leq \mathbb{E}\left(\int_0^{\eta} \mathbb{1}_{\{L(N^p s)\geq K\}} \right) + \eta\mathbb{P}(S_N \leq \eta),$$

again with the ergodic theorem and Relation (2.50), there exists some N_0 and $K>0$ such that $\mathbb{E}(\mu_N([0,\eta]\times[K,+\infty]))\leq\varepsilon$. Lemma 1.3 of Kurtz [55] shows that the sequence of random measures (μ_N) on $\mathbb{R}_+\times\mathbb{N}$ restricted to $[0,\eta]\times\mathbb{N}$ is tight.

From there and in the same way as in Section 2.C.1.2, it is not difficult to conclude the proof of the proposition, on $[0,\eta]$ and extend by induction this result on the time interval $[0,k\eta]$, for any $k\geq1$. □

Let $\mathcal{N}$ be a Poisson process on $\mathbb{R}_+^3$, independent of the Poisson processes $(\mathcal{P}_{Y,i})$, whose intensity measure is $\mathrm{d}s\otimes\mathrm{d}t\otimes\kappa_0\exp(-\kappa_0 a)\,\mathrm{d}a$. Recall that such a point process has the same distribution as

$$\sum_{n\geq0} \delta_{(s_n,t_n,E_n)},$$

where (s_n) and (t_n) are independent Poisson processes on $\mathbb{R}_+$ with rate 1 independent of the i.i.d. sequence (E_n) of exponential random variables with parameter κ_0. See Chapter 1 of [67].

Definition 2.18 (Time Change). *Define the process* $(A_N(t))$ *by*

$$A_N(t) \stackrel{\text{def.}}{=} \left(t + \int_{[0,t]\times\mathbb{R}_+} a\mathbb{1}_{\left\{Y_2^N(s-)=1\right\}} \mathcal{N}((0,\kappa_1 Y_1^N(s-)Y_2^N(s-)), \mathrm{d}s, \mathrm{d}a) \right),$$

and its associated inverse function as

$$B_N(t) \stackrel{\text{def.}}{=} \inf\left\{ s>0 : A_N(s) \geq t \right\}.$$

The instants of jump of $(A_N(t))$ correspond to the case when $(Y_2^N(t))$ could switch from 1 to 0 for the dynamic of $(X_2^N(t))$ and the size of the jump is the duration of time when $(X_2^N(t))$ stays at 0, its distribution is exponential with parameter κ_0.

The process $(A_N(t))$ gives in fact the correct timescale to construct the process $(X_N(t))$ with the process $(Y_N(t))$. We define the process $(\widetilde{X}_N(t))$ on $\mathbb{N}^2$ by, for $t\geq 0$,

$$\begin{cases} \widetilde{X}_N(A_N(t)) = Y_N(t), \\ \left(\widetilde{X}_1^N(u), \widetilde{X}_2^N(u)\right) = \left(Y_1^N(t-)-1, 0\right), u\in[A_N(t-), A_N(t)). \end{cases} \tag{2.51}$$

If t is a jump instant of $(A_N(t))$, the process does not change on the time interval $[A_N(t-), A_N(t))$. In this way, $(\widetilde{X}_N(t))$ is defined on $\mathbb{R}_+$.

LEMMA 2.19. *For $t>0$, then $A_N(B_N(t))=t$ if t is not in an interval $[A_N(u-), A_N(u))$ for some $u>0$, and the relation*

$$\sup_{t\geq 0} |\widetilde{X}_N(t) - Y_N(B_N(t))| \leq 1$$

holds.

PROOF. This is easily seen by an induction on the time intervals $[A_N(s_n), A_N(s_{n+1}))$, $n\geq 0$, where (s_n) is the sequence on instants of jump of $(A_N(t))$, with the convention that $s_0=0$. □

PROPOSITION 2.20. *The processes $(X_N(t))$ and $(\widetilde{X}_N(t))$ have the same distribution.*

PROOF. The proof is standard. The Markov property of $(\widetilde{X}_N(t))$ is a consequence of the Markov property of $(Y_N(t))$ and the strong Markov property of the Poisson process $\mathcal{N}$. It is easily checked that the Q-matrices of $(X_N(t))$ and $(\widetilde{X}_N(t))$ are the same. □

PROPOSITION 2.21. *For the convergence in distribution,*

$$\lim_{N\to+\infty} \left(\frac{A_N(t)}{N}\right) = (a(t)) \stackrel{\text{def.}}{=} \left(\alpha_1 \frac{1}{\kappa_0(e^{\kappa_2/\kappa_3}-1)} \left(1 - \exp\left(-\frac{\kappa_1\kappa_2}{\kappa_3} t\right)\right)\right). \tag{2.52}$$

PROOF. Let $T>0$. By using the fact that, for $0\leq u\leq T$, the relation

$$Y_1^N(u) \leq x_1^N + \mathcal{P}_{Y,1}((0,\kappa_0)\times(0,T])$$

holds, the sequence of processes

$$\left(\frac{1}{N}\int_0^t \frac{\kappa_1}{\kappa_0} \mathbb{1}_{\left\{Y_2^N(u)=1\right\}} Y_1^N(u)\,\mathrm{d}u\right)$$

is thus tight by the criterion of modulus of continuity. See Theorem 7.3 of Billingsley [14] for example. Proposition 2.17 shows that its limiting point is necessarily $(a(t))$.

We note that the process

$$(M_{A,N}(t)) = \left(\frac{1}{N}\left(A_N(t) - t - \frac{\kappa_1}{\kappa_0}\int_0^t \mathbb{1}_{\left\{Y_2^N(u)=1\right\}} Y_1^N(u)\,\mathrm{d}u\right)\right),$$

it is a square integrable martingale whose predictable increasing process is

$$(\langle M_{A,N}\rangle(t)) = \left(\frac{\kappa_1}{\kappa_0 N}\int_0^t \mathbb{1}_{\left\{Y_2^N(u)=1\right\}} \frac{Y_1^N(u)}{N}\,\mathrm{d}u\right).$$

The martingale is vanishing as N gets large by Doob's Inequality. The proposition is proved. □

Proposition 2.17 establishes a limit result for the sequence of processes $(Y_1^N(t)/N)$. In our construction of $(X_1^N(t))$, time intervals, whose lengths have an exponential distribution, are inserted. During these time intervals the coordinates do not change. To have a non-trivial limit result for $(X_1^N(t)/N)$, the timescale of the process has clearly to be sped-up. It turns out that the convenient timescale for this is (Nt), this is a consequence of the convergence in distribution of $(A_N(t)/N)$ established in Proposition 2.21.

PROPOSITION 2.22. *For the convergence in distribution, the relation*

$$\lim_{N\to+\infty}\left(B^N(Nt), t<t_\infty\right) = (a^{-1}(t)) = \left(-\frac{\kappa_3}{\kappa_1\kappa_2}\ln\left(\frac{\alpha_1-\kappa_0(e^{\kappa_2/\kappa_3}-1)t}{\alpha_1}\right), t<t_\infty\right), \tag{2.53}$$

holds, where $(a(t))$ is defined in Proposition 2.21 and

$$t_\infty = \frac{\alpha_1}{\kappa_0(\exp(\kappa_2/\kappa_3)-1)}.$$

PROOF. Note that both $(A_N(t))$ and $(B_N(t))$ are non-decreasing processes and that the relation $A_N(B_N(t))\geq t$ holds for all $t\geq 0$.

We are establishing the tightness property with the criterion of the modulus of continuity. The constants $\varepsilon>0$, $\eta>0$ are fixed. For $0<T<t_\infty$ we can choose $K>0$ sufficiently large so that $a(K)>T$ and we define

$$h_K = \inf_{s\leq K}\left(a(s+\eta)-a(s)\right),$$

clearly $h_K>0$. By definition of $(B_N(t))$, we have

$$\mathbb{P}(B_N(NT)\geq K) = \mathbb{P}\left(\frac{A_N(K)}{N}\leq T\right).$$

The convergence of Proposition 2.21 shows that there exists N_0 such that if $N\geq N_0$, the right-hand side of the last relation is less that ε and that

$$\mathbb{P}\left(\sup_{0\leq u\leq K}\left|\frac{A_N(u+\eta)-A_N(u)}{N}-(a(u+\eta)-a(u))\right|\geq\frac{h_K}{2}\right)\leq\varepsilon \tag{2.54}$$

holds.

For $\eta>0$, and $0\leq s\leq t\leq T$, if $B_N(Nt)-B_N(Ns)\geq\eta$ holds, then

$$A_N\left(B_N(Ns)+\eta\right)-A_N\left(B_N(Ns)\right)\leq N(t-s)$$

and if $\delta\leq h_K/4$, for $N\geq N_0$,

$$\mathbb{P}\left(\sup_{\substack{0\leq s\leq t\leq T\\ t-s\leq\delta_0}}|B_N(Nt)-B_N(Ns)|\geq\eta\right) \leq \varepsilon+\mathbb{P}\left(\inf_{0\leq u\leq K}\left(\frac{A_N\left(u+\eta\right)}{N}-\frac{A_N\left(u\right)}{N}\right)\leq\frac{h_K}{4}\right)\leq 2\varepsilon,$$

by Relation (2.54). The sequence of processes $(B_N(Nt))$ is therefore tight and any of its limiting points is a continuous process. The convergence of Proposition 2.21 shows that a limiting point has the same finite marginals as the right-hand side of Relation (2.53). The proposition is proved. □

THEOREM 2.23. *If $(X_N(t))=(X_1^N(t), X_2^N(t))$ is the Markov process associated to the CRN (2.37) whose initial state is $(x_1^N, b)\in\mathbb{N}^2$, $b\in\mathbb{N}$ and*

$$\lim_{N\to+\infty} x_1^N/N = \alpha_1>0,$$

then, the convergence in distribution

$$\lim_{N\to+\infty}\left(\frac{X_1^N(Nt)}{N}, t<t_\infty\right)=\left(1-\frac{t}{t_\infty}, t<t_\infty\right).$$

holds, with $t_\infty=\alpha_1/(\kappa_0(\exp(\kappa_2/\kappa_3)-1))$.

Proof. Proposition 2.17 and 2.22 show that the sequence of processes

$$\left(\left(\frac{Y_N(t)}{N}, t>0\right), (B_N(Nt), t<t_\infty)\right)$$

is converging in distribution to $((y_\infty(t)), (a^{-1}(t), t<t_\infty))$. Consequently, the relation

$$\lim_{N\to+\infty}\left(\frac{Y_N(B_N(Nt))}{N}, t<t_\infty\right)=\left(y_\infty\left(a^{-1}(t)\right), t<t_\infty\right)$$

holds for the convergence in distribution. We conclude the proof of the proposition by using Lemma 2.19. □

7.2.2. *Vertical Axis.* For $N\geq 1$, the initial state is $x_N(0)=(a, x_2^N)$, it is assumed that $a<p$ and

$$\lim_{N\to+\infty}\frac{x_2^N}{N}=1. \tag{2.55}$$

As seen in Section 7.1.2 when the first coordinate is strictly less than p, with a second coordinate of the order of N, it takes an amount of time of the order of N^{p-1} for the process $(X_1^N(t))$ to hit p. See Lemma 2.16. In a second, short phase, a decrease of the second coordinate takes place before returning below p. We now establish two limiting results

Lemma 2.24. *If* $(Z(z,N,t))$ *is the solution of the SDE* (2.44) *with initial state* $z<p$, *and* $T_Z(z,N)$ *is its hitting time of* p *then, the sequence* $(T_Z(z,N)/N^{p-1})$ *converges in distribution to an exponential random variable with parameter*

$$r_1\overset{\text{def.}}{=}\frac{\kappa_0}{(p-1)!}\left(\frac{\kappa_0}{\kappa_1}\right)^{p-1}. \tag{2.56}$$

Proof. The proof is standard. It can be done by induction on $p\geq 2$ with the help of the strong Markov property of $(Z(z,N,t))$ for example. □

We now study the phase during which $(X_1^N(t))$ is greater or equal to p. Define (T_k^N) the non-decreasing sequence of stopping time as follows, $T_0^N=0$ and, for $k\geq 0$,

$$T_{k+1}^N=\inf\{t\geq T_k^N : X_1^N(t)=p-1, X_1^N(t-)=p\}. \tag{2.57}$$

Proposition 2.25 (Decay of $(X_2^N(t))$). *Under Assumption 2.55 for the initial condition, for the convergence in distribution*

$$\lim_{N\to+\infty}\left(\frac{X_2^N(T_1^N)}{X_2^N(0)}, \frac{T_1^N}{X_2^N(0)^{p-1}}\right)\overset{\text{dist.}}{=}\left(U^{\delta_1}, E_1\right),$$

where U *is a uniform random variable on* $[0,1]$, *independent of* E_1 *an exponential random variable with parameter* r_1 *defined by Relation* (2.56), *and*

$$\delta_1\overset{\text{def.}}{=}\frac{\kappa_3(p-1)!}{\kappa_1}. \tag{2.58}$$

Proof. Let H_N be the hitting time of p for $(X_1^N(t))$, H_N has the same distribution as $T_Z(k, x_2^N)$. Its asymptotic behavior is given by Lemma 2.24. Since the reactions $pS_1+S_2\rightleftharpoons pS_1+2S_2$ are inactive on the time interval $[0, H_N]$, we have $X_2^N(H_N)=x_2^N+p-a\overset{\text{def.}}{=}x_2^\tau$. Let τ_N such that $H_N+\tau_N$ is the first instant when

$(X_1^N(t))$ returns to $p-1$. With the strong Markov property the time origin is translated to H_N, it is enough to study the asymptotic behavior of $X_2^N(\tau_N)$ starting from x_2^τ.

It is not difficult to see that, with high probability external arrivals do not play a role during the time interval $[0, \tau_N)$ simply because the other reaction rates are of the order of N or N^2. We will ignore them. We can therefore assume that $X_1^N(s) = p$ until τ_N. In the same way it is easily seen that the sequence of random variables $(N\tau_N)$ is tight.

After time 0, the transition $x \to x-e_2$ occurs until time $\tau_{1,N}$ when one of the other reactions happens. Since we are interested at the final value $X_2^N(\tau_{1,N})$, modulo a time change, it is equivalent to look at the Markov process with Q-matrix

$$x \longrightarrow x + \begin{cases} -e_1-e_2 & \kappa_1, \\ e_2 & \kappa_2(p-1)!, \\ -e_2 & \kappa_3(p-1)!(x_2-1)^+, \end{cases}$$

When $x \to x-e_1-e_2$ or $x \to x+e_2$ occurs, i.e. after an exponentially random variable F_1 with parameter $\kappa_1+\kappa_2(p-1)!$, the state of $(X_2(t))$ at this instant is

$$X_2^N(\tau_{1,N}-) \stackrel{\text{dist.}}{=} \sum_{i=1}^{x_2^\tau} \mathbb{1}_{\{E_i \geq F_1\}},$$

where (E_i) is an i.i.d. sequence of exponential random variables with parameter $\kappa_3(p-1)!$, and $\left|X_2^N(\tau_{1,N}) - X_2^N(\tau_{1,N}-)\right| \leq 1$. For the convergence in distribution,

$$\lim_{N \to +\infty} \frac{X_2^N(\tau_{1,N})}{X_2^N(0)} = \exp\left(-\kappa_3(p-1)!F_1\right).$$

The transition $x \to x-e_1-e_2$ occurs at time $\tau_{1,N}$ with probability $1-q_1$, with

$$q_1 = \frac{\kappa_2(p-1)!}{\kappa_1+\kappa_2(p-1)!},$$

and in this case $\tau_N = \tau_{1,N}$. Otherwise, there is a new cycle of length $\tau_{2,N}$ and that

$$\lim_{N \to +\infty} \frac{X_2^N(\tau_{1,N})+\tau_{2,N})}{X_2^N(0)} = \exp\left(-\kappa_3(p-1)!(F_1+F_2)\right),$$

where (F_i) is an i.i.d. sequence with the same distribution as F_1. By induction we obtain the convergence in distribution

$$\lim_{N \to +\infty} \frac{X_2^N(\tau_N)}{X_2^N(0)} = \exp\left(-\kappa_3(p-1)! \sum_1^G F_i\right),$$

where G is a random variable independent of (F_i) with a geometric distribution with parameter q_1, $\mathbb{P}(G \geq n) = q_1^{n-1}$ for $n \geq 1$. Trivial calculations gives the desired representation. □

In view of the last result it is natural to expect that the convergence of the scaled process $(X_2^N(t/N^{p-1})/N$ to a Markov process with jumps. The only problem is that, as we have seen in the last proof, the jumps downward of the limit process are due to a large number of small jumps downwards, of the order of N, on the time interval of length τ_N of the previous proof. Event if τ_N is arbitrarily small when N gets large, there cannot be convergence in the sense of the classical J_1-Skorohod topology. There are topologies on the space of càdlàg functions $\mathcal{D}(\mathbb{R}_+)$ for which convergence in distribution may hold in such a context. See Jakubowski [44] for example. For the sake of simplicity, we present a convergence result formulated for a weaker topology expressed in terms of occupation measures.

We now introduce a Markov process on $(0,1]$ as the plausible candidate for a limiting point of $(X_2^N(t/X_2^N(0)^{p-1})/N$.

DEFINITION 2.26. *The infinitesimal generator $\mathcal{A}$ of a Markov process $(U(t))$ on $(0,1]$ is defined by, for $f\in\mathcal{C}_c((0,1])$,*

$$\mathcal{A}(f)(x) = \frac{r_1}{x^{p-1}}\mathbb{E}\left(f\left(xV^{\delta_1}\right)-f(x)\right), \quad x\in(0,1], \tag{2.59}$$

where r_1, δ_1 are constants defined by Relations (2.56) and (2.58), and V is a uniform random variable on $[0,1]$.

Analytically the operator $\mathcal{A}$ can be expressed as

$$\mathcal{A}(f)(x) = \frac{r_1}{x^{p-1}}\int_0^1 \left(f\left(xu^{\delta_1}\right)-f(x)\right)\mathrm{d}u, \quad x\in(0,1].$$

PROPOSITION 2.27. *If $(U(t))$ is a Markov process on $(0,1]$ with infinitesimal generator $\mathcal{A}$, then, with probability 1, it is an explosive process converging to 0.*

PROOF. Assume that $U(0)=\alpha\in(0,1]$. By induction, the sequence of states visited by the process has the same distribution as (V_n) with, for $n\geq 0$,

$$V_n \stackrel{\text{def.}}{=} \alpha\exp\left(-\delta_1\sum_{i=1}^n E_i\right),$$

where (E_i) is an i.i.d. sequence of exponentially distributed random variables with parameter 1. The sequence of instants jumps has the same distribution as

$$\left(t_n^V\right) \stackrel{\text{def.}}{=} \left(\sum_{i=1}^n (V_{i-1})^{p-1}\frac{\phi_i}{r_1}\right),$$

where (ϕ_i) is an i.i.d. sequence of exponentially distributed random variables with parameter 1, independent of (E_i). It is easily seen that (V_n) converges to 0 almost surely and that the sequence $(\mathbb{E}(t_n^V))$ has a finite limit. The proposition is proved. □

DEFINITION 2.28 (Scaled occupation measure of $(X_2^N(t))$). *For $N\geq 1$, Λ_N is the random measure on $\mathbb{R}_+\times(0,1]$ defined by, for $f\in\mathcal{C}_c(\mathbb{R}_+\times(0,1])$,*

$$\langle\Lambda_N, f\rangle = \frac{1}{N^{p-1}}\int_0^{+\infty} f\left(\frac{s}{N^{p-1}}, \frac{X_2^N(s)}{N}\right)\mathrm{d}s. \tag{2.60}$$

We can now state our main scaling result for large initial states near the vertical axis.

THEOREM 2.29. *If $(X_N(t))$ is the Markov process associated to the CRN (2.37) whose initial state is $(a, x_2^N)\in\mathbb{N}^2$, $a\leq p-1$, and such that*

$$\lim_{N\to+\infty} x_2^N/N = \alpha > 0,$$

then the sequence (Λ_N) defined by Relation (2.60) converges in distribution to Λ, the occupation measure of $(U(t))$ a Markov process with infinitesimal generator $\mathcal{A}$ starting at α, i.e. for $f\in\mathcal{C}_c(\mathbb{R}_+\times(0,1])$,

$$\langle\Lambda, f\rangle = \int_0^{+\infty} f(s, U(s))\,\mathrm{d}s.$$

PROOF. Without loss of generality, due to the multiplicative properties of the convergence, see Proposition 2.25, we can take $\alpha=1$ and assume that $X_2^N(0)=N$. Recall that the Laplace transform of a random measure G on $\mathbb{R}_+\times(0,1]$ is given by

$$\mathcal{L}_G(f) \stackrel{\text{def.}}{=} \mathbb{E}\left(\exp\left(-\langle G, f\rangle\right)\right),$$

for a non-negative function $f \in \mathcal{C}_c(\mathbb{R}_+ \times (0,1])$. See Section 3 of Dawson [21].

To prove the convergence in distribution of (Λ_N) to Λ, it is enough to show that the convergence

$$\lim_{N \to +\infty} \mathcal{L}_{\Lambda_N}(f) = \mathcal{L}_{\Lambda}(f),$$

holds for all non-negative functions $f \in \mathcal{C}_c(\mathbb{R}_+ \times (0,1])$. See Theorem 3.2.6 of [21] for example.

If $f \in \mathcal{C}_c(\mathbb{R}_+ \times (0,1])$, its support is included in some $[0,T] \times (\eta, 1]$, for $\eta > 0$ and $T > 0$.

Let (T_k^N) the sequence of stopping times defined by Relation (2.57). The Laplace transform of Λ_N at f is given by

$$\mathcal{L}_{\Lambda_N}(f) = \mathbb{E}\left(\exp\left(-\sum_{k \geq 0} \int_{T_k^N/N^{p-1}}^{T_{k+1}^N/N^{p-1}} f\left(s, \frac{X_2^N\left(T_k^N\right)}{X_2^N(0)}\right) \mathrm{d}s\right)\right). \tag{2.61}$$

Let (t_k^V, V_k) be the sequence of couples of instants of jumps and its value of the Markov process $(V(t))$, as defined in the proof of Proposition 2.27. For $\varepsilon > 0$, there exists some n_0 such that

$$\mathbb{P}\left(\alpha \prod_{i=1}^{n_0} V_i \geq \frac{\eta}{2}\right) \leq \varepsilon/2,$$

holds, and, consequently,

$$\left|\mathcal{L}_{\Lambda}(f) - \mathbb{E}\left(\exp\left(-\sum_{k=0}^{n_0-1} \int_{t_k^V}^{t_{k+1}^V} f\left(s, V_k\right) \mathrm{d}s\right)\right)\right| \leq \varepsilon. \tag{2.62}$$

Proposition 2.25 shows that, for the convergence in distribution,

$$\lim_{N \to +\infty} \left(\frac{X_2^N(T_{k+1}^N)}{X_2^N(T_k^N)}, \frac{T_{k+1}^N - T_k^N}{X_2^N(T_k^N)^{p-1}}, k \geq 0\right) = \left(U_k^{\delta_1}, E_k, k \geq 0\right),$$

where (U_k) and (E_k) are i.i.d. independent sequence of random variables whose respective distributions are uniform on $[0,1]$, and exponential with parameter r_1. Hence, there exists N_0 such that if $N \geq N_0$, then

$$\begin{cases} \left|\mathcal{L}_{\Lambda_N}(f) - \mathbb{E}\left(\exp\left(-\sum_{k=0}^{n_0-1} \int_{T_k^N/N^{p-1}}^{T_{k+1}^N/N^{p-1}} f\left(s, \frac{X_2^N\left(T_k^N\right)}{X_2^N(0)}\right) \mathrm{d}s\right)\right)\right| \leq 2\varepsilon, \\ \mathbb{P}\left(\frac{X_2^N(T_{k+1}^N)}{X_2^N(T_k^N)} \leq 1, \forall k \in \{0, \ldots, n_0\}\right) \geq 1 - \varepsilon \end{cases} \tag{2.63}$$

Define, for $n > 0$,

$$(I_n^N) \stackrel{\text{def.}}{=} \left(\sum_{k=0}^{n-1} \int_{T_k^N/N^{p-1}}^{T_{k+1}^N/N^{p-1}} f\left(s, \frac{X_2^N\left(T_k^N\right)}{X_2^N(0)}\right) \mathrm{d}s\right)$$

In views of Relations (2.62) and (2.63), all we have to do is to prove, for every $n > 0$, the convergence in law of (I_n^N) to

$$I_n \stackrel{\text{def.}}{=} \int_0^{t_n^V} f\left(s, V(s)\right) \mathrm{d}s = \sum_{k=0}^{n-1} \int_{t_k^V}^{t_{k+1}^V} f\left(s, V_k\right) \mathrm{d}s,$$

as N gets large.

We will prove by induction on $n>0$, the convergence in distribution

$$\lim_{N\to+\infty}\left(I_n^N,\left|\ln\left(\frac{X_2^N\left(T_n^N\right)}{X_2^N(0)}\right)\right|,\frac{T_n^N}{X_2^N(0)^{p-1}}\right)$$
$$=\left(\int_0^{t_n^V} f\left(s,V(s)\right)\mathrm{d}s,\left|\ln\left(V_n\right)\right|,t_n^V\right).$$

We will show the convergence of the Laplace transform of the three random variables taken at (a,b,c), for a, b, $c>0$.

For $n=1$, this is direct consequence of Proposition 2.25. If it holds for $n\geq 1$, the strong Markov property of $(X^N(t))$ for the stopping time T_n^N gives the relation

$$H_N(a,b,c)\overset{\text{def.}}{=}\mathbb{E}\left(\exp\left(-aI_{n+1}^N-b\left|\ln\left(\frac{X_2^N\left(T_{n+1}^N\right)}{X_2^N(0)}\right)\right|-c\frac{T_{n+1}^N}{X_2^N(0)^{p-1}}\right)\middle|\mathcal{F}_{T_n^N}\right)$$
$$=\exp\left(-aI_n^N-b\left|\ln\left(\frac{X_2^N\left(T_n^N\right)}{X_2^N(0)}\right)\right|-c\frac{T_n^N}{X_2^N(0)^{p-1}}\right)$$
$$\times\Psi_N\left(\frac{X_2^N\left(T_n^N\right)}{X_2^N(0)},\frac{T_n^N}{X_2^N(0)^{p-1}}\right),$$

where, for $x>0$ and $u>0$, $\Psi_N\left(x,u\right)$ is defined as

$$\mathbb{E}_{(p-1,\lfloor Nx\rfloor)}\left(\exp\left(-a\int_0^{T_1^N/X_2^N(0)^{p-1}}f\left(s+u,x\right)\mathrm{d}s-b\left|\ln\left(\frac{X_2^N\left(T_1^N\right)}{X_2^N(0)}\right)\right|-c\frac{T_1^N}{X_2^N(0)^{p-1}}\right)\right).$$

Proposition 2.25, and the fact that the sequence $(N\tau_N)$ is tight in the proof of this proposition, gives the convergence

$$\lim_{N\to+\infty}\Psi_N\left(x,u\right)$$
$$=\mathbb{E}_x\left(\exp\left(-a\int_0^{E_{n+1}}f\left(s+u,x\right)\right)\mathrm{d}s-b\left|\ln\left(U_{n+1}^\delta\right)\right|-cE_{n+1}\right),$$

where U_{n+1} is a uniform random variable on $[0,1]$, independent of E_{n+1} an exponential random variable with parameter r_1. With the induction hypothesis for n, Lebesgue's Theorem and the strong Markov property of $(U(t))$, we obtain the convergence

$$\lim_{N\to+\infty}\mathbb{E}(H_N(a,b,c))=E\left[\exp\left(-aI_n-b\left|\ln V_n\right|-ct_n^V\right)\right.$$
$$\left.\times\exp\left(-a\int_{t_n^V}^{t_{n+1}^V}f\left(s,x\right)\mathrm{d}s-b\left|\ln\left(\frac{V_{n+1}}{V_n}\right)\right|-c\left(t_{n+1}^V-t_n^V\right)\right)\right]$$
$$=\mathbb{E}\left(\exp\left(-aI_{n+1}-b\left|\ln V_{n+1}\right|-ct_{n+1}^V\right)\right).$$

The theorem is proved. □

2.A. Classical Stability Results

Deterministic CRNs have been introduced in Section 2.6. We state the classical results of Feinberg [26] and Horn and Jackson [42] on the stability properties of deterministic CRNs. See Feinberg [27] for a broader picture of stability results for CRNs. See also Gunawardena [36] for a quick, comprehensive, overview of these results.

We briefly recall some definitions on CRNs.

(a) A *linkage class* is a connected component of the reaction graph. The quantity ℓ denotes the number of linkage classes.
(b) The CRN is *weakly reversible* if every connected component of its reaction graph is strongly connected.
(c) The *Stoichiometric space* S is the vector subspace of $\mathbb{R}^m$ generated by $y_r^+ - y_r^+$, $r \in \mathcal{R}$, its dimension is denoted by s.
(d) The deficiency δ of the CRN is $|C - \ell - s|$, where C is the total number of complexes.

The main theorem can now be stated.

THEOREM 2.30 (Feinberg (1979)). *Let $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ be a chemical reaction network with deterministic mass action kinetics, if it is weakly reversible and with zero deficiency, $\delta = 0$, then there is exactly one equilibrium for the dynamical system defined by Relation* (2.15)*. Furthermore, this equilibrium is locally asymptotically stable.*

Proposition 2.5 shows that a deterministic CRN can be seen as an asymptotic stochastic CRN provided that the reaction rates are scaled conveniently as in Relations (2.22). The interesting result of Anderson et al. [7] show that the invariant distribution of a class of stochastic CRNs can be expressed with the equilibrium of a deterministic CRN.

THEOREM 2.31. *Let $(X(t))$ be the Markov process associated to a stochastic chemical reaction network, irreducible on $\mathcal{E}_0 \subset N^n$, whose Q-matrix is given by Relation* (2.9) *then, under the assumptions of deficiency zero theorem, Theorem 2.30, if $c = (c_i)$ is the equilibrium point of the dynamical system* (2.15)*, then the positive measure on $\mathcal{E}_0$ defined by*

$$\pi(x) = \prod_{i=1}^{n} \frac{c_i^{x_i}}{x_i!} \tag{2.64}$$

is an invariant measure of $(X(t))$.

The proof consists in plugging a product formula in the invariant measure equations. It has been used to other classes of CRNs in subsequent works in Cappelletti and Wiuf [19] and Cappelletti and Joshi [18] and Jia et al. [45] and references therein. An interesting feature of this class of CRN is that the existence and uniqueness of an invariant distribution is essentially expressed in terms of the topological structure of the graph $\mathcal{R}$ and not on its associated weights (κ_r) as long as they are all positive.

2.B. Stability Results

The notations of Section 3 are used in this section. If F a subset of the state space $\mathcal{E}_0$, the hitting time of F is denoted as T_F,

$$T_F = \inf\{s \geq 0 : X(s) \in F\}.$$

The following proposition is a variation of Proposition 8.5 of Robert [67]. It is included for the sake of self-containedness.

PROPOSITION 2.32. *Let $(X(t))$ be an irreducible on $\mathcal{E}_0$, non-explosive, Markov process associated to a CRN network with Q-matrix defined by Relation* (2.8)*, if there exists a finite subset F of $\mathcal{E}_0$ such that $\mathbb{E}_x(T_F)$ is finite for all $x \in \mathcal{E}_0$, then $(X(t))$ is positive recurrent.*

PROOF. Let (s_n) be the sequence of stopping times defined by $s_1{=}t_1{+}T_F{\circ}\theta_{t_1}$ and, for $n{\geq}1$,

$$s_{n+1} = \inf\{u{\geq}s_n{+}t_1{\circ}\theta_{s_n} : X(u){\in}F\} = s_n{+}t_1{\circ}\theta_{s_n}{+}T_F \circ \theta_{s_n+t_1\circ\theta_{s_n}}.$$

and $M_n{=}X(s_n)$. It is easily seen that (M_n) is the sequence of successive visits to F of the embedded Markov chain $(X(t_n))$. In particular (M_n) is a Markov chain on the finite space F and it is irreducible since $(X(t_n))$ has this property. The Markov chain (M_n) is therefore positive recurrent, its unique invariant probability distribution is denoted as π_F. Define the positive measure π on $\mathcal{E}_0$ by, if f is a non-negative function on $\mathcal{E}_0$,

$$\langle \pi, f\rangle = \int_{\mathcal{E}_0} f(x)\pi(\mathrm{d}x) = \sum_{x\in F} \mathbb{E}_x\left(\int_0^{s_1} f(X(u))\,\mathrm{d}u\right)\pi_F(x).$$

Its mass is

$$\pi(\mathcal{E}_0) = \sum_{x\in F} \mathbb{E}_x\left(s_1\right)\pi_F(x) = \sum_{x\in F} \pi_F(x)\left(\mathbb{E}_x\left(t_1\right){+}\mathbb{E}_{X(t_1)}\left(T_F\right)\right),$$

since there is only a finite number of possibilities for the first jump of $(X(t))$ starting from x, the mass is finite. We can therefore assume that, up to a scaling factor, π is a probability distribution on $\mathcal{E}_0$.

For $t{\geq}0$, we now show that if $X(0)\overset{\text{dist.}}{=}\pi$, then $X(t)\overset{\text{dist.}}{=}\pi$. If f is a non-negative function on $\mathcal{E}_0$,

$$\begin{aligned}
\mathbb{E}_\pi(f(X(t))) &= \sum_{x\in\mathcal{E}_0}\pi(x)\mathbb{E}_x(f(X(t))) = \sum_{x\in\mathcal{E}_0}\mathbb{E}_{\pi_F}\left(\int_0^{s_1}\mathbb{1}_{\{X(u)=x\}}\,\mathrm{d}u\right)\mathbb{E}_x(f(X(t))) \\
&= \sum_{x\in\mathcal{E}_0}\int_{u=0}^{+\infty}\mathbb{P}_{\pi_F}\left(X(u){=}x, u<s_1\right)\mathbb{E}_x(f(X(t)))\,\mathrm{d}u \\
&= \int_{u=0}^{+\infty}\mathbb{E}_{\pi_F}\left(f(X(u+t))\mathbb{1}_{\{u<s_1\}}\right)\mathrm{d}u,
\end{aligned}$$

with the Markov property and Fubini's theorem, hence,

$$\mathbb{E}_\pi(f(X(t))) = \mathbb{E}_{\pi_F}\left(\int_t^{t+s_1} f(X(u))\,\mathrm{d}u\right).$$

We note that, with the strong Markov property,

$$\begin{aligned}
&\mathbb{E}_{\pi_F}\left(\int_{s_1}^{t+s_1} f(X(u))\,\mathrm{d}u\right) \\
&\qquad = \mathbb{E}_{\pi_F}\left(\mathbb{E}_{X(s_1)}\left(\int_0^t f(X(u))\,\mathrm{d}u\right)\right) = \mathbb{E}_{\pi_F}\left(\int_0^t f(X(u))\,\mathrm{d}u\right),
\end{aligned}$$

since the distribution of $X(s_1){=}M_1$ is π_F. By gathering these results we finally obtain that $\mathbb{E}_\pi(f(X(t))) = \langle\pi, f\rangle$. The proposition is proved. □

2.C. Technical Proofs

The proofs of this section, although not difficult, are detailed for the sake of completeness, and also to show that some ingredients of a scaling analysis are elementary. For basic results on martingale theory and classical stochastic calculus, see Rogers and Williams [69].

To investigate scaling properties of stochastic CRNs, we use the formulation in terms of stochastic differential equations (SDE)introduced in Relation (2.10) of Section 2 to describe the Markov process.

2.C.1. Example of Section 5.1. There is only one complex of size 2 and one chemical reaction with a quadratic reaction rate. The associated Markov process $(X^N(t))=(X_1^N(t), X_2^N(t))$ can be represented as a solution of the following SDEs,

$$\begin{cases} \mathrm{d}X_1^N(t) &= \mathcal{P}_2((0,\kappa_2 X_2^N(t-)),\mathrm{d}t)-\mathcal{P}_1((0,\kappa_1 X_1^N(t-)),\mathrm{d}t) \\ \mathrm{d}X_2^N(t) &= \mathcal{P}_1((0,\kappa_1 X_1^N(t-)),\mathrm{d}t)-\mathcal{P}_{12}((0,\kappa_{12} X_1^N(t-)X_2^N(t-)),\mathrm{d}t), \end{cases} \tag{2.65}$$

with $X^N(0)=(x_1^N, x_2^N)$. See Relation (2.10).

We prove the convergence results of of Proposition 2.11. Note that the first convergence (2.29) is a direct consequence of Proposition 2.10.

2.C.1.1. *Proof of the convergence* (2.30) *in Proposition 2.11.* The initial state is such that

$$\lim_{N\to+\infty}\left(\frac{x_1^N}{\sqrt{N}}, \frac{x_2^N}{N}\right) = (\beta, 1)$$

with $\beta\geq 0$, the scaled process is defined in this case as

$$\left(\overline{X}^N(t)\right) \stackrel{\text{def.}}{=} \left(\frac{X_1^N(t/\sqrt{N})}{\sqrt{N}}, \frac{X_2^N(t/\sqrt{N})}{N}\right).$$

The integration of Relations of SDE (2.65) gives

$$\begin{cases} \overline{X}_1^N(t) &= \dfrac{x_1^N}{\sqrt{N}}+M_1^N(t)+\kappa_2\displaystyle\int_0^t \overline{X}_2^N(s)\,\mathrm{d}s-\frac{\kappa_1}{\sqrt{N}}\int_0^t \overline{X}_1^N(s)\,\mathrm{d}s \\ \overline{X}_2^N(t) &= \dfrac{x_2^N}{N}+M_2^N(t)+\dfrac{\kappa_1}{N}\displaystyle\int_0^t \overline{X}_1^N(s)\,\mathrm{d}s-\kappa_{12}\int_0^t \overline{X}_1^N(s)\overline{X}_2^N(s)\,\mathrm{d}s, \end{cases} \tag{2.66}$$

where, for i=1, 2, $(M_i^N(t))$ is a martingale whose increasing process is given by

$$\begin{cases} \left(\left\langle M_1^N\right\rangle(t)\right) &= \dfrac{\kappa_2}{\sqrt{N}}\displaystyle\int_0^t \overline{X}_2^N(s)\,\mathrm{d}s+\frac{\kappa_1}{N}\int_0^t \overline{X}_1^N(s)\,\mathrm{d}s \\ \left(\left\langle M_2^N\right\rangle(t)\right) &= \dfrac{\kappa_1}{N^2}\displaystyle\int_0^t \overline{X}_1^N(s)\,\mathrm{d}s+\frac{\kappa_{12}}{N}\int_0^t \overline{X}_1^N(s)\overline{X}_2^N(s)\,\mathrm{d}s. \end{cases} \tag{2.67}$$

From there, the method is quite simple, Relation (2.66) and Gronwall's lemma give the relation, for $T>0$ and $i\in\{1,2\}$,

$$\sup_N \sup_{t\leq T} \mathbb{E}\left(\overline{X}_i^N(t)\right) < +\infty.$$

With Relation (2.67) and Doob's Inequality, we obtain that the relation

$$\lim_{N\to+\infty} \mathbb{E}\left(\sup_{t\leq T}|M_i^N(t)|\right) = 0$$

holds. Using again Relation (2.66) and Gronwall's lemma, we get

$$\sup_N \mathbb{E}\left(\sup_{t\leq T}\overline{X}_i^N(t)\right) < +\infty, \qquad i\in\{1,2\}.$$

The next step uses the criterion of the modulus of continuity, see Theorem 7.3 of Billingsley [14]. If $(X(t))$ is some càdlàg process, $\delta>0$ and $T\geq 0$,

$$\omega_X(\delta) = \sup\{|X(t)-X(s)| : s,t\leq T, |t-s|\leq\delta\},$$

then Relation (2.66) gives

$$\omega_{\overline{X}_1^N}(\delta) \leq 2\sup_{t\leq T}|M_1^N(t)|+\kappa_2\delta\sup_{t\leq T}\overline{X}_2^N(t)+\frac{\kappa_1}{\sqrt{N}}\delta\sup_{t\leq T}\overline{X}_1^N(t). \tag{2.68}$$

For ε and $\eta>0$, we can fix $N_0\in\mathbb{N}$ and $K>0$ such that, for $N\geq N_0$ and $i\in\{1,2\}$,

$$\mathbb{P}\left(\sup_{t\leq T}|M_1^N(t)|>\eta/2\right)\leq\varepsilon, \qquad \mathbb{P}\left(\sup_{t\leq T}\overline{X}_i^N(t)\geq K\right)\leq\varepsilon.$$

If δ is chosen as $\min(\eta\sqrt{N_0}/(\kappa_1 K), \eta/(\kappa_2 K))$, by using Relation (2.68), we have, for all $N \geq N_0$,

$$\mathbb{P}\left(\omega_{\overline{X}_1^N}(\delta) > 3\eta\right) \leq 3\varepsilon.$$

The sequence $(\overline{X}_1^N(t))$ is tight. With the same argument, it is easily seen that $(\overline{X}_2^N(t))$ is also tight since the process $(X_1^N(t)X_2^N(t), t \leq T)$ is upper bounded by K^2 with high probability. The sequence of processes $(\overline{X}^N(t)$ is therefore tight. Relation (2.66) gives that any of its limiting point satisfies Relation (2.31). Convergence (2.30) is proved.

2.C.1.2. *Proof of the convergence* (2.32) *in Proposition 2.11.*
The initial state is $x^N = (x_1^N, k)$ for some $k \in \mathbb{N}$, and x_1^N such that

$$\lim_{N \to +\infty} \frac{x_1^N}{N} = 1.$$

A glance at the equation of X_2^N shows that, if X_1^N is the order of N, then the process $(X_2^N(t))$ is an $M/M/\infty$ CRN on a fast timescale. To handle such a setting, we follow the general method, see Kurtz [55], to establish an averaging principle.

Define, for $i = 1, 2$,

$$\overline{X}_i^N(t) \stackrel{\text{def.}}{=} \frac{X_i^N(t)}{N}.$$

In the same way as in the proof of the convergence (2.30) in Proposition 2.11 in Section 2.C.1.1, Gronwall's Lemma, it is not difficult to show that the sequence of processes $(\overline{X}_1^N(t))$ is tight for the convergence in distribution.

We have, for $t \geq 0$,

$$X_1^N(t) \geq \sum_{i=1}^{X_1^N(0)} \mathbb{1}_{\left\{E_{\kappa_1, i} \geq t\right\}},$$

where $(E_{\kappa_1, i})$ is an i.i.d. sequence of exponential random variables with parameter κ_1. Consequently and using proposition 2.9, for any $\varepsilon > 0$ and $T > 0$, there exist $K_0 > 0$, $\eta > 0$ and N_0 such that, if $N \geq N_0$,

$$\mathbb{P}(\mathcal{E}_N) \geq 1 - \varepsilon, \text{ with } \mathcal{E}_N = \left\{\eta \leq \inf_{t \leq T} \overline{X}_1^N(t), \sup_{t \leq T} \overline{X}_1^N(t) \leq K_0\right\}. \tag{2.69}$$

On the event $\mathcal{E}_N$, one can construct a coupling such that, for $t \leq T$, $X_2^N(t) \leq L(Nt)$, where $(L(t))$ is the process associated to an $M/M/\infty$ queue with input rate $\kappa_1 K_0$ and output rate $\kappa_2 \eta$.

(a) *Limit of* $(\overline{X}_1(t))$*:* The integration of the first relation of the SDE (2.65) gives the identity

$$\overline{X}_1^N(t) = \frac{x_1^N}{N} + M_1^N(t) + \kappa_2 \int_0^t \overline{X}_2^N(s)\,\mathrm{d}s - \kappa_1\sqrt{N} \int_0^t \overline{X}_1^N(s)\,\mathrm{d}s, \tag{2.70}$$

where $(M_1^N(t))$ is a martingale whose increasing process is given by

$$\left(\left\langle M_1^N \right\rangle(t)\right) = \frac{1}{N}\left(\kappa_2 \int_0^t \overline{X}_2^N(s)\,\mathrm{d}s + \kappa_1 \int_0^t \overline{X}_1^N(s)\,\mathrm{d}s\right).$$

The upper bound of $(X_2^N(t))$, Lebesgue's Lemma and the tightness of $(\overline{X}_1^N(t))$ gives the convergence in distribution to (0) of $(\left\langle M_1^N \right\rangle(t))$ and therefore, with Doob's Inequality, of the martingale $(M_1^N(t))$. The upper

bound of $(X_2^N(t))$ and Lebesgue's Theorem show that the sequence of processes

$$\left(\int_0^t \overline{X}_2^N(s)\,\mathrm{d}s\right)$$

converges in distribution to 0. It is then easy to conclude the proof of convergence (2.30) in Proposition 2.11.

(b) *Convergence of the occupation measure of* $(X_2^N(t))$*:* The upper bound of (X_2^N) is enough to show the convergence of $(\overline{X}_1^N(t))$, but we can also have a convergence result on the species S_2, looking at its *occupation measure*. Define (μ_N) the sequence of random measures defined by, for $N\geq 1$ and if f is a continuous function with compact support on $\mathbb{R}_+\times\mathbb{N}$, $f\in\mathcal{C}_c(\mathbb{R}_+\times\mathbb{N})$,

$$\langle \mu_N, f\rangle = \int_{\mathbb{R}_+\times N} f(s, X_2^N(s))\,\mathrm{d}s.$$

To prove the tightness of (μ_N), it is enough to show that, for $\varepsilon'>0$, there exists $C>0$ such that $\mathbb{E}(\mu_N([0,T]\times[0,C]))\geq(1-\varepsilon')T$. See Lemma 1.3 of Kurtz [55]. For $N\geq N_0$ sufficiently large we have for all $C>0$,

(2.71)
$$\mathbb{E}(\mu_N([0,T]\times[C,+\infty])) \leq \int_0^T \mathbb{P}(X_2^N(s)\geq C)\,\mathrm{d}s \leq \varepsilon T + \int_0^T \mathbb{P}(L(Ns)\geq C)\,\mathrm{d}s.$$

Since, for $s>0$, $(L(Ns))$ converges in distribution to $L(\infty)$ a random variable with a Poisson distribution, with Lebesgue's Theorem we obtain the relation

$$\lim_{N\to+\infty}\int_0^T \mathbb{P}(L(Ns)\geq C)\,\mathrm{d}s = T\mathbb{P}(L(\infty)\geq C),$$

and this last term can be made arbitrarily small for some large C. The sequence (μ_N) is therefore tight. We can take a subsequence (N_p) such that $((\overline{X}_1^{N_p}(t)), \mu_{N_p})$ is converging in distribution to $((x_1(t)), \mu_\infty)$. Lemma 1.4 of Kurtz [55] shows that μ_∞ can be represented as

$$\langle \mu_\infty, f\rangle = \int_{\mathbb{R}_+\times N} f(s,x)\pi_s(\mathrm{d}x)\,\mathrm{d}s,$$

where (π_t) is an optional process with value in the space of probability distribution on $\mathbb{N}$. Note that, almost surely, $x_1(t)=x_{c,1}(t)=\exp(-\kappa_1 t)>0$ for all $t\leq T$.

Let g be a function on $\mathbb{N}$ with finite support, then the second relation of SDE (2.65) gives

$$g(X_2^{N_p}(t)) = g(k)+M_f^{N_p}(t)+\kappa_1\int_0^t X_1^{N_p}(s))\nabla^+(g)(X_2^{N_p}(s))\,\mathrm{d}s + \kappa_{12}\int_0^t X_1^{N_p}(s)X_2^{N_p}(s)\nabla^-(g)(X_2^{N_p}(s))\,\mathrm{d}s,$$

where $\nabla^\pm(f)(x)=f(x\pm 1)-f(x)$, $x\geq 0$. From there, by dividing both terms of this relation by N_p and let p go to infinity, with standard arguments, the martingale $(M_f^{N_p}(t))$ vanishes and we obtain the relation

$$\int_0^t\int_{\mathbb{N}} \left(\kappa_1 x_1(s)\nabla^+(g)(x)+\kappa_{12}x_1(s)x\nabla^-(g)(x)\right)\pi_s(\mathrm{d}x)\,\mathrm{d}s = 0,$$

or

$$\int_0^t x_1(s) \int_{\mathbb{N}} \Omega(g)(x)\pi_s(\mathrm{d}x)\,\mathrm{d}s = 0,$$

where Ω is the infinitesimal generator of an $M/M/\infty$ CRN with input rate κ_1 and output rate κ_{12}. By using the fact that a.s. $(x_1(t))$ is positive and continuous and the càdlàg property of (π_t), we obtain that, almost surely, for all $t{\leq}T$,

$$\int_{\mathbb{N}} \Omega(g)(x)\pi_t(\mathrm{d}x) = 0,$$

for all indicator functions of a finite subset of $\mathbb{N}$. Therefore, almost surely for all $t{\leq}$, π_t is a Poisson distribution on $\mathbb{N}$ with parameter κ_1/κ_{12}.

2.D. General Triangular Topologies

In this section, we show that every triangular CRN with general complexes is stable.

Definition 2.33. *A* triangular network *is a CRN* $(\mathcal{S},\mathcal{C},\mathcal{R})$ *with*

$$\begin{cases} \mathcal{S} &= \{S_1,\dots,S_m\} \\ \mathcal{C} &= \{y_1,y_2,y_3\}, \\ \mathcal{R} &= \{(y_1,y_2),(y_2,y_3),(y_3,y_1)\}, \end{cases} \tag{2.72}$$

where y_ℓ, $\ell{\in}\{1,2,3\}$ are distinct elements of $\mathbb{N}^m$. The reaction constants are $(\kappa_i, 1{\leq}i{\leq}3)$.

The CRN is represented by the following graph,

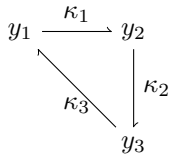


Figure 3. Triangular CRN

We are going to show the following proposition:

Proposition 2.34. *The Markov process associated to a triangular network is positive recurrent.*

By using the deficiency zero theorem, we show that it is enough to investigate a smaller class of CRNs. After this simplification, we simply have to prove the positive recurrence property of a one-dimensional Markov process on $\mathbb{N}$, which can be done using the Foster-Lyapunov criterion, see Corollary 2.17.

2.D.1. Reduction. Let $(\mathcal{S},\mathcal{C},\mathcal{R})$ be a triangular network, the space $\mathbb{S}$ is its stoichiometric subspace, see Section 2.A of the Appendix,

$$\mathbb{S} \overset{\text{def.}}{=} \mathrm{Span}\{y_2{-}y_1, y_3{-}y_2, y_1{-}y_3\}.$$

If $(X(t))$ is its associated Markov process, with initial state x, it is straightforward to show that the process is irreducible on $(x{+}\mathbb{S}) \cap \mathbb{N}^m$,

When $\dim \mathbb{S}{=}2$, the CRN is clearly weakly reversible with deficiency 0, the stochastic version of the deficiency zero theorem, Theorem 2.31 gives that the process is positive recurrent.

We are left with the CRN with $\dim \mathbb{S}=1$. Consequently, we can find some k_1, k_2, $k_3 \in \mathbb{Z}^*$ such that

$$k_1(y_2-y_1) = k_2(y_3-y_2) = k_3(y_1-y_3) \tag{2.73}$$

Since the three complexes differ, we know that $y_i - y_{i-1} \neq 0$.

First, we are going to consider the processes that have a kind of mass conservation property.

LEMMA 2.35. *If there exists $\rho \in (0,+\infty)^m$ such that for all $t \geq 0$,*

$$\langle X(t), \rho \rangle \overset{\text{def.}}{=} \sum_{i=1}^{m} \rho_i X_i(t) = \langle x, \rho \rangle ,$$

then $(X(t))$ is positive recurrent.

PROOF. The process is irreducible in a finite state space, it is positive recurrent. □

Let $u=y_2-y_1$. If we can find coordinates $1 \leq i,j \leq m$ such that $u_i<0$ and $u_j>0$, then we can find some $\rho \in (0,+\infty)^m$ such that $\langle u, \rho \rangle = 0$. Relation (2.73) shows that the process $(\langle X(t), \rho \rangle)$ is not changed by its jumps.

We can assume that our CRN is such that $u \in \mathbb{R}_+^m$ or $u \in \mathbb{R}_-^m$. If we have $u_i=0$ for some $i \in \{1,\dots,m\}$, then $y_{1,i} = y_{2,i} = y_{3,i}$ and the number of elements of the species i is kept constant all along the reactions, therefore $X_i(t) = x_i$ for all $t \geq 0$. Up to a change in the rates of the CRN with a fixed parameter, we can remove the i-th coordinate of the system, without changing the sample paths of the others components of the process. We can now suppose that $u_i \neq 0$ for all $1 \leq i \leq m$.

The complexes can be rewritten as y_1, y_1+u and $y_1+(1+k_2/k_1)u$. Every reaction adds or remove an amount of "u". The evolution of the process is therefore expressed in the"number of u" in the system. Since k_2/k_1 is not necessarily an integer, in order to have an integer valued process, we proceed as follows. With some arithmetic manipulations, we can find a $\Delta \in (\mathbb{N} \setminus \{0\})^m$ and p_1, p_2, $p_3 \in \mathbb{Z}^*$ such that

$$y_2 - y_1 = p_1\Delta, \quad y_3 - y_2 = p_2\Delta \quad \text{and} \quad y_1 - y_3 = p_3\Delta,$$

with $p_1+p_2+p_3=0$. The quantity Δ is the "smallest positive vector" on the line $\mathbb{S} \cap (\mathbb{N}^*)^m$, and it is invariant by a rotation of indexes of complexes. Up to a rotation of indexes, we can assume that $p_1>0$ and $p_1+p_2>0$, by taking y_1 as the complex with the smallest norm for example.

Now we are left with triangular networks with three complexes of the form y_1, $y_1+p_1\Delta$ and $y_1+(p_1+p_2)\Delta$.

Using Proposition 2.38, we can get rid of the y_1 in the three complexes. We only have to show the positive recurrence of the process associated to the CRN of Figure 4, with p_1, $(p_1+p_2) \geq 1$.

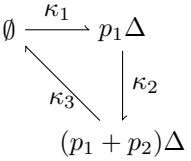


FIGURE 4. Reduced triangular CRN

2.D.2. A one dimensional process. From now on, $(Y(t))$ is the Markov process associated to the CRN of Figure 4, with initial state y. The process $(Y(t))$ is living in the subset $H(y)=\{y+k\Delta, k\in\mathbb{Z}\}\cap\mathbb{N}^d$, a "line" H_a in $\mathbb{N}^d$,

$$H_a \stackrel{\text{def.}}{=} \{a+k\Delta, k\geq 0\}, \quad \text{with} \quad a = y - \left\lfloor \frac{\langle y,\Delta\rangle}{\langle\Delta,\Delta\rangle} \right\rfloor \Delta.$$

The quantity a is the vector of $H(y)$ with the smallest norm. The process $(Y(t))$ lives in H_a, and every state y' of H_a is entirely determined by the number of Δ contained in $y'-a$. We define Φ_a the function on H_a by

$$\Phi_a(z) = \frac{\langle z-a,\Delta\rangle}{\langle\Delta,\Delta\rangle}$$

and its inverse function on $\mathbb{N}$,

$$\Psi_a(k) = a+k\Delta.$$

Setting $(N(t)) = (\Phi_a(Y(t)))$, we have $(\Psi_a(N(t))=(Y(t))$ and, using the SDE associated with $(Y(t))$, we can show that $(N(t))$ is a continuous time Markov jump process on $\mathbb{N}$, with infinitesimal generator $\mathcal{Q}$ defined by, for $k\in\mathbb{N}$,

$$\begin{aligned}(2.74)\quad \mathcal{Q}(f)(k) = \kappa_1(f(k+p_1)-f(k)) + \kappa_2\left(\Psi_a(k)\right)^{(p_1\Delta)}(f(k+p_2)-f(k)) \\ + \kappa_3\left(\Psi_a(k)\right)^{((p_1+p_2)\Delta)}(f(k-p_1-p_2)-f(k)),\end{aligned}$$

for any function f with finite support on $\mathbb{N}$. This process still has polynomial rates in k, since, for $y\in\mathbb{N}^m$,

$$\left(\Psi_a(k)\right)^{(y)} = \prod_{i=1}^{m} \frac{(a_i+k\Delta_i)!}{(a_i+k\Delta_i-y_i)!}.$$

LEMMA 2.36. *The process $(N(t))$ defined by Equation (2.74) is positive recurrent.*

PROOF. Taking r the greatest common divisor of p_1 and p_2, then $(N(t))$ is clearly irreducible on

$$r\mathbb{N} \stackrel{\text{def.}}{=} \{rk, \quad k\geq 0\}.$$

To show the positive recurrence of $(N(t))$ we use Corollary 2.4 with the energy function $f_0(k)=k$.

Let $q_1=\max\{p_1,p_1+p_2\}$, $q_2=\min\{p_1,p_1+p_2\}$ and $\kappa_{\min}=\min\{\kappa_1,\kappa_2,\kappa_3\}$. Since $q_1-q_2\geq 1$ and $q_1\geq 1$, we can show the following inequality:

$$\limsup_{k\to+\infty} \frac{\mathcal{Q}(f_0)(k)}{\left(\Psi_a(k)\right)^{(q_1\Delta)}} \leq -\frac{\kappa_{\min}}{2},$$

and since $\left(\Psi_a(k)\right)^{(q_1\Delta)}$ goes to infinity when k goes to infinity, there is some $K_0>0$ such that, for $k\geq K_0$,

$$\mathcal{Q}(f_0)(k) \leq -1,$$

which proves the positive recurrence of $(N(t))$. □

Consequently, $(Y(t))=(\Psi_a(N(t))$ is a positive recurrent Markov process. Proposition 2.34 is proved.

COROLLARY 2.37. *The Markov process associated to a weakly reversible CRN with three complexes is positive recurrent.*

PROOF. First, note that if we add some reactions to our triangular CRN, our proof can still be used. The only CRNs remaining are the ones with the following graph:

$$y_1 \rightleftharpoons y_2 \rightleftharpoons y_3.$$

Using the same arguments as in our proof, we can reduce the set of CRNs to the ones with a graph of the type

$$\emptyset \rightleftharpoons p_1\Delta \rightleftharpoons p_2\Delta \quad \text{or} \quad p_1\Delta \rightleftharpoons \emptyset \rightleftharpoons p_2\Delta,$$

for some $\Delta \in (\mathbb{N} - \{0\})^m$ and $p_1, p_2 \geq 1$, $p_1 \neq p_2$. Here, we can use the same argument to reduce our process to a process on $\mathbb{N}$, which can be shown to be positive recurrent using the same arguments as in the proof of Proposition 2.34. □

This proof can be done for three complexes, but the generalization to a CRN with four or more complexes seems unlikely, at least with our approach.

PROPOSITION 2.38. *Let $(\mathcal{S} = \{1, \dots, m\}, \mathcal{C}, \mathcal{R})$ be a CRN, $(X(t))$ its associated continuous time Markov jump process starting at $x \in \mathbb{N}^m$. For some $i \in \{1, \dots m\}$, suppose that for every $y \in \mathcal{C}$, $y_i \geq 1$. We define the CRN $(\mathcal{S}', \mathcal{C}', \mathcal{R}')$ as*

$$\begin{cases} \mathcal{S}' = \mathcal{S} \\ \mathcal{C}' = \{y - e_i, \quad y \in \mathcal{C}\} \\ \mathcal{R}' = \{(y^- - e_i) \rightharpoonup (y^+ - e_i),\ y^- \rightharpoonup y^+ \in \mathcal{R}\}. \end{cases}$$

where $e_i = (\delta_{j=i})_{1 \leq j \leq m}$. Let $(Y(t))$ be the process associated to $(\mathcal{S}', \mathcal{C}', \mathcal{R}')$ starting from $y = x - e_i$. If $x_i \geq 1$, $(X(t))$ is positive recurrent on $H_x \subset \mathbb{N}^m$ if and only if $(Y(t))$ is positive recurrent on $H_x - e_i = \{z - e_i,\ z \in H_x\}$.

PROOF. The proof is straightforward, seeing that the embedded Markov chains of both processes have the same transition rates. Note that if $x_i < 1$, none of the reactions can happen and the process $(X(t))$ is constant equal to x, therefore positive recurrent. □

CHAPTER 3

Stochastic Chemical Reaction Networks with Discontinuous Limits and AIMD processes

Contents

1. Introduction

The chemical reaction network (CRN) with m chemical species considered in this paper is represented as

$$\begin{aligned}(3.1)\quad \emptyset \xrightarrow{\kappa_0} S_1 \xrightarrow{\kappa_1} S_1+S_2 \xrightarrow{\kappa_2} \cdots \\ \xrightarrow{\kappa_i} S_i+S_{i+1} \xrightarrow{\kappa_{i+1}} S_{i+1}+S_{i+2} \xrightarrow{\kappa_{i+2}} \cdots \\ \xrightarrow{\kappa_{m-1}} S_{m-1}+S_m \xrightarrow{\kappa_m} S_m \xrightarrow{\kappa_{m+1}} \emptyset.\end{aligned}$$

The associated Markov process $(X(t))=(X_i(t), 1\leq i\leq m)$ lives in the state space $\mathbb{N}^m$. The kinetics considered for these networks is the classical *law of mass action*. See Guldberg and Waage [35] and Voit et al. [76] for example. It has the Q-matrix $Q=(q(x,y), x, y\in\mathbb{N}^m)$ defined by, for $x\in\mathbb{N}^m$ and $2\leq i<m-1$,

$$(3.2)\qquad \begin{cases} q(x, x+e_{i+1}-e_{i-1}) = \kappa_i x_i x_{i-1}, \qquad q(x, x+e_1) = \kappa_0, \\ q(x, x+e_2) = \kappa_1 x_1, \qquad q(x, x-e_{m-1}) = \kappa_m x_m x_{m-1}, \\ q(x, x-e_m) = \kappa_{m+1} x_m, \end{cases}$$

where $(\kappa_i)\in(0,+\infty)^{m+2}$ is the vector of the reaction rates.

Note that the molecules of chemical species S_1 are created from an external input and the chemical species S_m vanish independently of the other chemical species. See Section 2 for a detailed presentation of the mathematical context of these CRNs.

We have two important features of this class of CRNs.

(a) *Quadratic Rates.*
Due to the assumption of the law of mass action, the rate of most of reactions of Relation (3.1) is a quadratic function of the state. It is a polynomial function in general.

The polynomial dependence for the reaction rates has the consequence that different timescales may coexist in the dynamical behavior of CRNs. In our example, if all coordinates are of the order of N, the rate of the

reaction in the middle of Relation (3.1) is $O(N^2)$, for the reaction $S_m \rightharpoonup \emptyset$, it is $O(N)$ and for $\emptyset \rightharpoonup S_1$, it is only $O(1)$.

(b) *Boundary Behaviors.*
The reaction in the middle of Relation (3.1) does not occur if either x_i or x_{i+1} is null. A molecule of S_i may be transformed into a molecule of S_{i+2} only if there is at least a molecule of S_{i+1}, even if the $(i+1)$th coordinates is not changed by the reaction. This is a boundary effect, some reactions do not occur on the boundary of the state space.

These boundary effects may have a strong impact on the time evolution of CRNs. For example, for the CRN of Laurence and Robert [52],

$$\emptyset \rightleftharpoons S_1+S_2, \qquad pS_1+S_2 \rightleftharpoons pS_1+2S_2,$$

it has been shown, Theorem 32, that if $p \geq 2$ to investigate the time evolution of the CRN starting from large initial states of the order of N, one has to speed-up time by a factor N^{p-1}, i.e. consider limit theorems with the fast timescale $t \to N^{p-1}t$, to get non-trivial time evolution. This is essentially due to the boundary condition that at least p copies of S_1 are required in two reactions.

As we will see the boundary behaviors of the CRNs of this paper have also an impact on its time evolution, through a phenomenon referred to as *discrete-induced transitions in the literature*. See Section 1.2.

Auto-catalytic Reaction Networks.
There are other classical classes of CRNs with a generic related way of transforming chemical species. The reactions for these networks are

$$S_i+S_j \rightharpoonup 2S_i, \qquad \emptyset \rightleftharpoons S_i, \qquad 1 \leq i \neq j \leq m.$$

See Togashi and Kaneko [74] for example. As it will be seen, they seem to share some of the properties of the CRNs we are considering in this paper.

1.1. Scaling with the Norm of the Initial State. We will investigate this class of CRNs via the convergence in distribution of its scaled sample paths. This may provide an interesting insight on the time evolution of these CRNs. Note that the deficiency of these CRNs is 1, see Section 2, the classical result of Anderson et al. [7] cannot be used here, so that even the existence of an invariant distribution is not known a priori.

Classically, a scaling approach is mainly done via a scaling of reaction rates. This can be achieved in several ways.

(a) The reaction rates are scaled so that all chemical reactions have a rate of the same order of magnitude in N, then with a convenient scaling of the space variable, it can be shown that the scaled Markov process converges in distribution to a deterministic dynamical system of the type described above. See Mozgunov et al. [60] and Section 2.3 of Laurence and Robert [52]. The drawback of this scaling is that the scaled CRN does not really exhibit anymore different timescales in N since the transitions of the scaled process are all $O(1)$.

(b) The reference Ball et al. [11] considers several CRNs with given scaled of reaction rates with some parameter N, a multi-timescale analysis of several classes of such CRNs is achieved via proofs of averaging principles. See also Kang and Kurtz [46] and Kang et al. [47].

In the spirit of Laurence and Robert [52], the scaling we consider in this paper does not change the basic dynamic of the CRN, in particular its reaction rates. It is assumed that the initial state of the CRN is "large", its norm is proportional to some scaling parameter N. We investigate the time evolution of the CRN, in

particular how such a saturated initial state returns to some neighborhood of the origin.

If the initial states of two chemical species S_i, S_{i+1} are both of the order of N, the rate of the reaction in the middle of Relation (3.1) is of the order of N^2 which is maximal for this class of CRNs. Proposition 3.2 of Section 3 shows that, for any $\varepsilon>0$, the state of the CRN goes "quickly" to a set S_N of states for which the indices of the coordinates whose value is greater than εN are at distance at least two, the other coordinates being $o(N)$.

When the number of nodes is odd and the initial state is such that the coordinates with an odd index are of the order of N and all the others are $O(1)$, Theorem 3.5 shows that $(X_N(t)/N)$ is converging in distribution to the solution $(x(t))$ of an ODE converging to 0 at infinity. The decay of the state of the CRN starting from this initial state is therefore observed on the normal time scale. This is one of the few general results, with respect to m, we have been able to derive. Nevertheless it turns out that small values of m provide already non-trivial behaviors.

In Section 4, the case of the network with three nodes (chemical species), m=3, is considered. It is shown, see Proposition 3.6, that for a set of initial states the process $(X_N(t)/N)$ converges in distribution to a continuous, but random, process. The stochastic fluctuations, represented by the martingales in the evolution equations vanish, as usual, due to the scaling procedure. Nevertheless there remains a random, discrete, component in the limit. This is due a boundary behavior of the kinetics. This case provides an example of a CRN whose first order is not the solution of a set of deterministic ODEs.

1.2. A Scaling Picture with Discontinuous Stochastic Processes. A CRN with four chemical species is investigated in Sections 6 and 7. A class of initial states gives rise to a more complex behavior than what we have observed when m=3. We did not try a complete (cumbersome) classification of initial states from this point of view as it has been done for m=3, but we do believe that this is *the* interesting class of initial states.

Recall that N is the scaling parameter of the initial state. The initial states considered are of the type $(0, N, 0, 0)$, with the convention that "0", resp. "N", means $O(1)$, resp. $O(N)$. We show that the process lives in the subset of the state space of elements of the type $(0, N, 0, \sqrt{N})$ and that the decay of the norm of the state occurs on the timescale $(\sqrt{N}t)$. More important, this decay is in fact based on a Discrete-Induced Transitions phenomenon (DIT) which we now describe.

$$\emptyset \xrightarrow{\kappa_0} S_1 \xrightarrow{\kappa_1} S_1{+}S_2 \xrightarrow{\kappa_2} S_2{+}S_3 \xrightarrow{\kappa_3} S_3{+}S_4 \xrightarrow{\kappa_4} S_4 \xrightarrow{\kappa_5} \emptyset$$

FIGURE 1. CRN with 4 chemical species

Discrete-Induced Transitions (DIT). One of the early references to this phenomenon is Togashi and Kaneko [74] where the term has been coined apparently. In the context of auto-catalytic CRNs (see above) with few nodes, it is observed, via numerical experiments, that a limited number of molecules of one chemical species can switch the entire bio-chemical state of a system. It is characterized by the fact that the state variable alternates between two subsets of the state spaces. From a biological point of view, the role of this set of molecules can be seen as a "switch" which can block or activate a set of chemical reactions. See also Togashi and Kaneko [73] and Saito and Kaneko [71]. The mathematical characterization

of this phenomenon has mainly been done with the analysis of the associated invariant measure. Intuitively the DIT phenomenon should be expressed by the fact that the invariant distribution is concentrated on at least two subsets of the state space, a bi-modal distribution. See Bibbona et al. [13], Hoessly and Mazza [41] and Gallinger and Popovic [33]. Note however that it does not give a dynamical picture of this phenomenon, such as an estimation of the sojourn time in a subset of states before visiting another subset. This is associated to the metastability property of statistical physics. See Bovier and Hollander [15]. A related phenomenon has also been analyzed in del Sole [22].

Returning to our CRN with four chemical species and the initial state of the type $(0, N, 0, \sqrt{N})$, it turns out that the growth of the fourth coordinate $(X_4(t))$ occurs only during time intervals whose duration are $O(1/\sqrt{N})$ and during them there is a large number of positive jumps of this process, of the order of $\sqrt{N}$. Recall that this phenomenon is only due to the law of mass action which drives the kinetics of the CRN. This is where boundary effects have a significant impact.

The switch effect occurs during these small time intervals, the occurrence of them is driven by the isolated instants of creation of particles of chemical species S_3. The duration of these time intervals vanishes in the limit, so we do not have a metastability-related phenomenon as discussed in the literature. Nevertheless they play a critical role in the kinetics of the system, since this is at these instants, and only there, that the second coordinate $(X_2(t))$ can decrease.

AIMD Processes. The asymptotic behavior of $(X_4(t)/\sqrt{N})$ can be described in terms of jump processes belonging to a class of AIMD processes on $\mathbb{R}_+$. These processes exhibit an exponential decay between jumps or a negative jump proportional to the current value of the state (Multiplicative Decrease) and, for the increasing part, positive jumps depending on the current state or a linear growth (Additive Increase). There are two classes of AIMD processes in our analysis, their infinitesimal generators Ω_0 and Ω_1 are given by, for bounded function $f \in \mathcal{C}_1(\mathbb{R}_+)$ and $v \geq 0$,

$$\Omega_0(f)(v) = -\frac{1}{\gamma} v f'(v) + \int_0^{+\infty} \left(f\left(\sqrt{v^2 + 2\beta s}\right) - f(v) \right) e^{-s} \, \mathrm{d}s, \tag{3.3}$$

$$\Omega_1(f)(v) = \frac{1}{\gamma} f'(v) + v \int_0^1 \left(f\left(v u^{\beta}\right) - f\left(v\right) \right) \mathrm{d}u, \tag{3.4}$$

for some constants β, $\gamma > 0$.

AIMD processes play also an important role in several classes of stochastic models, in mathematical finance, see Bertoin and Yor [12] and Yor [80] for an overview, communication networks Guillemin et al. [34], or genomics Cowan and Chiu [20]. An analogous multiplicative property has also been observed in other CRNs. See Section 8 of Laurence and Robert [52] for example. On the normal timescale, the main convergence result is given by the following theorem.

THEOREM. *If $X_N(0) = (0, N, 1, v_N)$, with $(v_N/\sqrt{N})$ converging to $v \geq 0$, then for the convergence in distribution for the M_1-Skorohod topology and also the S-topology,*

$$\lim_{N \to +\infty} \left(\frac{X_2^N(t)}{N}, \frac{X_4^N(t)}{\sqrt{N}} \right) = (1, V(\kappa_0 t)),$$

where $(V(t))$ is the Markov process on $\mathbb{R}_+$ whose infinitesimal generator Ω_0 is given by, for $f \in \mathcal{C}_c^1(\mathbb{R}_+)$ and $x \in \mathbb{R}_+$,

$$\Omega_0(f)(x) = -\frac{\kappa_5}{\kappa_0} x f'(x) + \int_0^{+\infty} \left(f\left(\sqrt{x^2 + 2\frac{\kappa_3}{\kappa_4} s}\right) - f(x) \right) e^{-s} \, \mathrm{d}s$$

The limit of the scaled process is not only random, as in Section 4.1 for m=3, but also discontinuous. The fact that the limit is a jump process is in fact a consequence of the large number of positive jumps during small time intervals. The M_1-Skorohod topology and the S-topology are weaker than the classical J_1-Skorohod topology which does not allow accumulation of jumps. See the discussion in Section 6.

The convergence result of $(X_2(t)/N)$ to (1) implies that the normal timescale is too slow to observe a decay of the norm. See Laurence and Robert [52] for a discussion of this phenomenon. We prove the following convergence in distribution, it involves a speed-up of the timescale by a factor of $\sqrt{N}$.

THEOREM. *If $X(0)=(0,N,0,v_N)$, with $(v_N/\sqrt{N})$ converging to $v\geq 0$, then the relation*

$$\lim_{N\to+\infty}\left(\frac{X_2\left(\sqrt{N}t\right)}{N}, t<t_\infty\right)=\left(\left(1-\frac{t}{t_\infty}\right)^2, t<t_\infty\right)$$

holds for the convergence in distribution, with

$$t_\infty=\sqrt{2}\,\frac{\kappa_5}{\kappa_0^2}\sqrt{\frac{\kappa_4}{\kappa_3}}\,\Gamma\left(\frac{\kappa_0}{2\kappa_5}\right)\Big/\,\Gamma\left(\frac{\kappa_5+\kappa_0}{2\kappa_0}\right),$$

where Γ is the classical Gamma *function.*

See Theorem 3.24. The proof of this result relies on several ingredients:

(a) The proof of a limit result, Theorem 3.12, related to AIMD processes associated to an infinitesimal generator of the type Ω_1 of Relation (3.4). An explicit expression of their invariant distributions, Section 5, is used;
(b) Multiple time-changes. This is an interesting example of the efficiency of stochastic calculus in such a context;
(c) The proof of a stochastic averaging principle.

We also show that the fourth coordinate of a related process is of the order of $\sqrt{N}$, with a limit result for its associated occupation measure, and the values of the first and third coordinates are essentially 0 or 1 in the limit, from the point of view of their contribution in the evolution equations.

1.3. Outline of the Paper. Section 2 introduces the class of CRNs investigated, the associated stochastic models and the kinetic equations. Section 3 proves a scaling result for a CRN with an odd number of nodes and a specific class of initial states. A scaling analysis of the CRN with three nodes is done in Section 4. Section 5 introduces and investigates the AIMD processes of interest for our scaling analysis. Sections 6 and 7 are devoted to the CRN with four nodes and a class of initial states.

2. Stochastic Model

We first recall the formal definitions for the model of CRN of interest.

DEFINITION 3.1. *A* chemical reaction network *(CRN) with m* chemical species*, $m\geq 1$, is defined by a triple $(\mathcal{S},\mathcal{C},\mathcal{R})$,*

— *$\mathcal{S}=\{1,\ldots,m\}$ is the set of species;*
— *the set of* complexes*, is a finite subset of $\mathbb{N}^m$*

$$\mathcal{C}=\{0,\{e_1\},\{e_m\},\{e_i+e_{i+1}\},1\leq i\leq m-1\},$$

where e_i is the ith unit vector of $\mathbb{N}^m$.

— *The set of* chemical reactions $\mathcal{R}$*, is a subset of* $\mathcal{C}^2$*,*

$$\mathcal{R} = \{(e_i{+}e_{i+1}, e_{i+1}{+}e_{i+2}), 1{\le}i{<}m{-}2\} \cup \{(0,e_1),(e_1,e_1{+}e_2),(e_m{+}e_{m-1},e_m),(e_m,0)\}$$

The notation 0 refers to the complex associated to the null vector of $\mathbb{N}^m$, $\emptyset{=}(0)$. A chemical reaction $r{=}(y_r^-, y_r^+){\in}\mathcal{R}$ corresponds to the change of state, for $x{=}(x_i)$ and $y_r^{\pm}{=}(y_{r,i}^{\pm})$,

$$x \longrightarrow x{+}y_r^+{-}y_r^-,$$

provided that $y_{r,i}^-{\le}x_i$ holds for $1{\le}i{\le}m$, i.e. there are at least $y_{r,i}^-$ copies of chemical species of type i, for all $i{\in}\mathcal{S}$, otherwise the reaction cannot happen. For the CRNs considered here, we have $y_{r,i}^-{\in}\{0,1\}$ for all $r{\in}\mathcal{R}$ and $1{\le}i{\le}m$.

The Markov process $(X(t))$ is clearly irreducible on $\mathbb{N}^m$. Its Q-matrix associated to the law of mass action is given by Relation (3.2). For $2{\le}i{<}m{-}1$, the transitions of $(X(t))$ transform chemical species S_{i-1} to S_{i+1} via the action of S_i. As a particle system, the particles arrive at rate κ_0 at node 1 or at node 2 at rate $\kappa_1 x_1$ and move on the right by steps of size 2 to finally leave the network at either node $m{-}1$ or m. The boundary effect is that the transformation of a molecule of S_{i-1} occurs at a rate proportional to the number of molecules of S_i. The parity of the indices of the nodes plays an important role as it can be expected.

It is easily seen that this CRN is *weakly reversible* and its *deficiency* is 1, see Feinberg [27] for the definitions.

2.1. Definitions and Notations. It is assumed that on the probability space there are $m{+}2$ independent Poisson processes on $\mathbb{R}_+^2$, $\mathcal{P}_i$, $0{\le}i{\le}m{+}1$, with intensity $\mathrm{d}s{\otimes}\,\mathrm{d}u$ and also an independent i.i.d. family of Poisson process $\mathcal{N}_\sigma$, $\sigma{>}0$, on $\mathbb{R}_+^2{\times}\mathbb{R}_+{\times}\mathbb{R}_+^{\mathbb{N}}$ whose intensity measure is

$$\mathrm{d}s{\otimes}\,\mathrm{d}t{\otimes}\sigma\exp(-\sigma a)\,\mathrm{d}a{\otimes}Q(\mathrm{d}b),$$

where Q is the distribution on $\mathbb{R}_+^{\mathbb{N}}$ of an i.i.d. sequence (E_i) of exponential random variables with parameter κ_5.

The underlying filtration used throughout the paper is $(\mathcal{F}_t)$, with, for $t{\ge}0$,

$$\mathcal{F}_t = \sigma\left(\mathcal{P}_i(A{\times}[0,s)), 0{\le}i{\le}m{+}1, \mathcal{N}_\sigma(A{\times}[0,s){\times}B{\times}C), \sigma{>}0, A,B{\in}\mathcal{B}(\mathbb{R}_+), C{\in}\mathcal{B}(\mathbb{R}_+^{\mathbb{N}})s{\le}t\right).$$

All stopping time and martingale properties implicitly refer to this filtration. See Section B.1 of Laurence and Robert [52]. If H is a metric space, $\mathcal{C}_c(H)$ denotes the set of continuous functions on H with compact support and $\mathcal{C}_c^1(\mathbb{R}_+)$ the subset of $\mathcal{C}_c(\mathbb{R}_+)$ of continuously differentiable functions on $\mathbb{R}_+$. For $T{>}0$ $\mathcal{D}([0,T),\mathbb{R})$ denotes the set of càdlàg functions on $[0,T)$, that is right continuous functions with left limits at all point.

For $\rho{>}0$, we will denote by $\mathrm{Pois}(\rho)$ the Poisson distribution on $\mathbb{N}$ with parameter ρ. Due to the numerous processes that have to be considered and to avoid heavy notations, we will use in the text the same notations such as $(A_N(t))$, $(B_N(t))$, $(Z_N(t))$ for different stochastic processes, or $(M_N(t))$ for a martingale, in different contexts, essentially in proofs of results. Similarly, several stopping times are denoted as τ_N with possibly other indices.

2.2. Stochastic Differential Equations. The goal of this paper is of investigating the transient properties of the sample paths of these CRNs and in particular to describe, via a functional limit theorem, how the process $(X(t))$ starting from a

"large" initial state comes back to a neighborhood of the origin. It is assumed that the sequence of initial states satisfies the relation,

$$\lim_{N\to+\infty}\left(\frac{X_i^N(0)}{N}\right)=\alpha=(\alpha_i)\in\mathbb{R}_+^m. \tag{3.5}$$

The scaling parameter N used is such that the norm of the initial state is of the order of N.

For $N\geq 1$, the Markov process $(X_N(t))=(X_i^N(t))$ starting from $X_N(0)$ with Q-matrix Q can be represented as the solution of the SDEs, for $1<i<m$,

$$\begin{cases}
\mathrm{d}X_1^N(t) &= \mathcal{P}_0\left((0,\kappa_0),\mathrm{d}t\right)-\mathcal{P}_2\left(\left(0,\kappa_2X_1^N(t-)X_2^N(t-)\right),\mathrm{d}t\right),\\
\mathrm{d}X_2^N(t) &= \mathcal{P}_1\left(\left(0,\kappa_1X_1^N(t-)\right),\mathrm{d}t\right)-\mathcal{P}_3\left(\left(0,\kappa_3X_2^N(t-)X_3^N(t-)\right),\mathrm{d}t\right),\\
\mathrm{d}X_i^N(t) &= \mathcal{P}_{i-1}\left(\left(0,\kappa_{i-1}X_{i-2}^N(t-)X_{i-1}^N(t-)\right),\mathrm{d}t\right)\\
&\qquad -\mathcal{P}_{i+1}\left(\left(0,\kappa_{i+1}X_i^N(t-)X_{i+1}^N(t-)\right),\mathrm{d}t\right),\\
\mathrm{d}X_m^N(t) &= \mathcal{P}_{m-1}\left(\left(0,\kappa_{m-1}X_{m-2}^N(t-)X_{m-1}^N(t-)\right),\mathrm{d}t\right)\\
&\qquad -\mathcal{P}_{m+1}\left(\left(0,\kappa_{m+1}X_m^N(t-)\right),\mathrm{d}t\right).
\end{cases} \tag{3.6}$$

See Rogers and Williams [69] for example.

The scaled process is introduced as

$$\left(\overline{X}_N(t)\right)=\left(\overline{X}_i^N(t)\right)=\left(\frac{X_i^N(t)}{N}\right). \tag{3.7}$$

2.3. The M/M/∞ queue. We finish by recalling the definition of a process associated to a very simple, but important, CRN,

$$\emptyset \underset{\mu}{\overset{\lambda}{\rightleftharpoons}} S_1.$$

This is the $M/M/\infty$ queue with input parameter $\lambda\geq 0$ and output parameter $\mu>0$. It is a Markov process $(L(t))$ on $\mathbb{N}$ with transition rates

$$x\longrightarrow\begin{cases}x+1 & \lambda\\ x-1 & \mu x.\end{cases}$$

The invariant distribution of $(L(t))$ is Poisson with parameter $\rho=\lambda/\mu$. For $t\geq 0$, if $L(0)=0$, then $L(t)$ has a Poisson distribution with parameter

$$\rho(1-e^{-\mu t}).$$

If T_N is the hitting time of N, $T_N=\inf\{t>0:L(t)\geq N\}$, then the sequence

$$\left(\frac{\rho^N}{(N-1)!}T_N\right)$$

converges in distribution to an exponential random variable. As a consequence, for $p\geq 1$, the convergence in distribution

$$\lim_{N\to+\infty}\left(\frac{L(N^pt)}{N}\right)=(0)$$

holds. See Chapter 6 of Robert [67]. This is an important process in the context of stochastic CRNs, see Laurence and Robert [52].

3. Scaling Properties

In this section, we investigate some general properties of the asymptotic behavior of the sample paths of $(X(t))$ when the initial state is "large".

As it will be seen, this is a challenging problem in general. Sections 6 and 7 investigate a specific class of initial states of a network with *four* nodes for which a scaling description of the return path to 0 involves, at the normal timescale, *jump processes* and not the nice solution of some set of ODES as it is usually the case.

The scaling parameter N can be thought, up to some fixed multiplicative constant, as the norm of the initial state. Our first result of this section, Proposition 3.2, shows that, for any $\varepsilon>0$, the process goes "quickly" to a set of states for which the indices of the coordinates whose value is greater than εN are at distance at least two, the other coordinates being $o(N)$.

The second result, Theorem 3.5, considers the case when the number of nodes m of the CRN is odd and the initial state is of the order of $(\alpha_1, 0, \alpha_3, 0, \ldots, 0, \alpha_m)N$, with $\alpha_k>0$, for $k\in\{1, 3, \ldots, m\}$. It is shown that, on the normal timescale, the scaled process converges in distribution to the solution of a system of ODEs. An averaging principle is proved to establish this convergence. Its proof uses several results on the Markov process associated to a series of $M/M/\infty$ queues, Proposition 3.3 together with a coupling result, Proposition 3.4 for the proof of the tightness of occupation measures.

The following proposition states essentially that, from this point of view, one can concentrate the study on asymptotic initial states such that the positive components have isolated indices.

PROPOSITION 3.2. *Under Condition* (3.5), *if* $H_0 \stackrel{\text{def.}}{=} \left\{x\in\mathbb{R}_+^m : x_i x_{i+1} = 0, \forall 1\le i<m\right\}$, *then there exists* $\beta\in H_0$ *such that, for any* $\eta>0$, *there exists a stopping time* τ_N *satisfying the relation*

$$\lim_{N\to+\infty} \mathbb{P}\left(\max_{1\le i\le m}\left(\left|\frac{X_i^N(\tau_N)}{N} - \beta_i\right|\right) \ge \eta\right) = 0,$$

and the sequence (τ_N) *is converging in distribution to* 0.

In Section 4, there is an analogous result for the CRN with three nodes. We show that there is a convenient stopping time, possibly depending on N, when the coordinates are arbitrarily close to the set of states with the "correct" orders of magnitude.

PROOF. Define the scaled process

$$\left(\widetilde{X}_N(t)\right) = \left(\frac{X_N(t/N)}{N}\right).$$

Proposition 13 of Laurence and Robert [52] gives that the sequence of processes $(\widetilde{X}_N(t))$ is converging in distribution to $(x(t))=(x_i(t))$ the solution of the system of ODEs, for $2<i<m$,

$$\begin{cases} \dot{x_1}(t) = -\kappa_2 x_1(t) x_2(t),\ \dot{x_2}(t) = -\kappa_3 x_2(t) x_3(t), \\ \dot{x_i}(t) = \kappa_{i-1} x_{i-2}(t) x_{i-1}(t) - \kappa_{i+1} x_i(t) x_{i+1}(t), \\ \dot{x_m}(t) = \kappa_{m-1} x_{m-2}(t) x_{m-1}(t). \end{cases} \tag{3.8}$$

We now show that $(x(t))$ converges to $\beta\in H_0$ as t gets large.

If at $t_0>0$, $\max_i |x_i(t_0)-\beta_i|<\eta/2$, taking $\tau_N=t_0/N$, we have the desired result. For the convergence of $(x(t))$ to $\beta\in H_0$, we proceed by induction on the number of species. When $\alpha_1=0$, from Relations (3.8) we have $(x_1(t))\equiv(0)$. Otherwise, if $\alpha_1>0$ and $\alpha_2=0$, then, clearly, $(x_1(t))\equiv(\alpha_1)$ and $(x_2(t))\equiv(0)$.

It is therefore enough to consider the case $\alpha_1>0$ and $\alpha_2>0$. We have that $x_k(t)>0$ for all $t>0$ and $k\in\{1,\dots,m\}$. Indeed, this is true for $k=1$, 2 and then, by induction on $k\in\{3,\dots,m\}$, for all the other indices. Similarly, for all $2\le k<m$, the functions

$$\left(\sum_{i\ge 0:2i+1\le k} x_{2i+1}(t)\right) \text{ and } \left(\sum_{i\ge 1:2i\le k} x_{2i}(t)\right)$$

are (strictly) decreasing. Again by induction on $1\le i<m$, the function $(x_i(t))$ has therefore a limit β_i when t goes to infinity and we also obtain the convergence of $(x_m(t))$ when t goes to infinity, to a limit $\beta_m>0$. Relations (3.8) give the identity $\beta_i\beta_{i+1}=0$ for all $1\le i<m$. The proposition is proved. □

PROPOSITION 3.3. *For $p\ge 1$, let $(Y_i(t))$ be a Markov process on $\mathbb{N}^p$ with Q-matrix $R_{\lambda,\mu}=(r_{\lambda,\mu}(y,z))$ is such that, for $y=(y_i)\in\mathbb{N}^p$ and $1\le i\le p$,*

$$\begin{cases} r_{\lambda,\mu}(y,y+e_1)=\lambda, \quad r_{\lambda,\mu}(y,y-e_p)=\mu_p y_p,\\ r_{\lambda,\mu}(y,y+e_{i+1}-e_i)=\mu_i y_i, \end{cases}$$

where e_i is the ith unit vector of $\mathbb{N}^p$ and $\mu=(\mu_i, i=1,\dots,p)$ and λ are positive constants.

(a) *The invariant distribution of $(Y(t))$ is the product of p Poisson distributions with respective parameters λ/μ_i;*

(b) *If, for $N\ge 1$, $(Y_i^N(t))$ is the Markov process with Q-matrix R and initial state (y_i^N) such that the sequence (y_i^N/N) converges to (0),*

— *For any $\varepsilon>0$,*

$$\sup_{N\ge 1}\sup_{t>\varepsilon}\mathbb{E}\left(\sum_{i=1}^{p} Y_i^N(Nt)\right)<+\infty; \tag{3.9}$$

— *The relation*

$$\lim_{N\to+\infty}\left(\frac{Y^N(Nt)}{N}\right)=(0)$$

holds for the convergence in distribution.

PROOF. The statement on the invariant distribution is straightforward to prove.

The Markov process can be described as a kind of Ehrenfest model with p urns. The distribution of the sojourn time of a particle in the ith urn, $1\le i\le p$ is exponential with parameter μ_i, after which it moves to urn $i+1$. The process $(Y_i^N(t))$ can be expressed as the sum of two independent processes, $(Y_{i,0}^N(t)+Y_{i,1}^N(t))$, where $Y_{i,0}^N(t)$, resp. $Y_{i,1}^N(t)$ is the number of initial particles, resp. new particles (i.e. arrived after time 0), present in the ith urn at time t. The initial state of $(Y_{i,1}^N(t))$ is (0) in particular.

For $1\le i\le p$, we denote by $(E_{i,k}^0, k\ge 1)$ and $(E_{i,k}^0, k\ge 1)$ i.i.d. independent sequences of exponential random variables with parameter μ_i. The process of the total number of initial particles present in the system has the same distribution as

$$\left(\sum_{i=1}^{p}\sum_{k=1}^{y_i^N}\mathbb{1}_{\left\{E_{i,k}^0+E_{i+1,k}^0+\cdots+E_{p,k}^0\ge t\right\}}\right).$$

The two assertions of 2) of the proposition for $(Y_{i,0}^N(t)$ are easily checked.

If (t_n) is a Poisson process on $\mathbb{R}_+$ with parameter λ, then the arrivals of particles at the ith urn has the same distribution as $(t_n+E_{1,n}+E_{2,n}+\cdots+E_{i-1,n})$, i.e. a Poisson process with parameter λ. Consequently, the process $(Y_{i,1}^N(t))$ has the same distribution as the process of an $M/M/\infty$ queue of Section 2.3, starting empty with

arrival rate λ and service rate μ_i. The two assertions of 2) are a consequence of the properties of this model recalled in Section 2.3.

□

PROPOSITION 3.4 (Coupling). *If $\kappa_0{=}0$ and $m{=}2p{+}1$ and under Condition (3.5), if $\alpha{\in}H_0$ is such that $\alpha_{2i}{=}0$ for all $1{\le}i{\le}p$ and $\alpha_{2i+1}{>}0$, for all $0{\le}i{\le}p$, then there exist $\eta{>}0$ and a coupling of $(X_N(t)){=}(X_k^N(t), 1{\le}k{\le}m)$ and $(Y_i(Nt), 1{\le}i{\le}p)$, where $(Y(t))$ is a Markov process with Q-matrix $R_{\lambda,\mu}$ defined in Proposition 3.3, for some $\lambda{>}0$, with $\mu{=}(\mu_i){=}\kappa_{2i+1}\eta$, such that the relation*

$$\sum_{k=1}^{p} X_{2k}^N(t) \le \sum_{k=1}^{p} Y_k^N(Nt) \tag{3.10}$$

holds for all $t{\le}T_N$, with

$$T_N = \inf\left\{t{>}0 : \exists i{\in}\{1,\dots,p\}, X_{2i+1}^N(t){<}\eta N\right\}. \tag{3.11}$$

PROOF. We fix $\eta{=}\min\{\alpha_{2i+1}/2 : 1{\le}i{\le}p\}$ and $\lambda{>}0$ such that the relation

$$X_1^N(0){+}X_3^N(0){+}\cdots{+}X_{2p+1}^N(0) \le \lambda N$$

holds for all $N{\ge}1$. The condition $\kappa_0{=}0$ implies that $X_1^N(t){+}X_3^N(t){+}\cdots{+}X_{2p+1}^N(t)$ is also bounded by λN for all $t{>}0$.

We can represent the process $(Y_N(t)){=}(Y_i^N(t))$ as the solution of the following SDEs, for $2{\le}i{\le}p$:

$$\begin{cases} \mathrm{d}Y_1^N(t) &= \mathcal{P}_1((0,\kappa_1\lambda N),\mathrm{d}t) - \mathcal{P}_3((0,\kappa_3\eta N Y_1^N(t{-})),\mathrm{d}t), \\ \mathrm{d}Y_i^N(t) &= \mathcal{P}_{2i-1}((0,\kappa_{2i-1}\eta N Y_{i-1}^N(t{-})\eta N),\mathrm{d}t) \\ &\qquad\qquad -\mathcal{P}_{2i+1}((0,\kappa_{2i+1}\eta N Y_i^N(t{-})),\mathrm{d}t). \end{cases}$$

with $(Y_N(0)){=}(X_{2i}^N(0)){=}(x_{2i}^N)$, where the $\mathcal{P}^i$ are introduced in the SDE (3.6).

For convenience, the two processes are described in terms of a queueing system, they are respectively referred to as the X-system for $(X_{2i}^N(t), 1{\le}i{\le}2p{+}1)$ and the Y-system, for the process $(Y_i(t), 1{\le}i{\le}p)$. Initially there are x_{2i}^N customers in the ith queue for both systems.

External customers enter the system, at the rate $\kappa_1 X_1^N(t)$ in the X-system, and $\kappa_1\lambda N$ in the Y-system, therefore more customers enter the Y-system than the X-system. Furthermore, since we choose η such that $X_{2i+1}(t){\ge}\eta N$ for all $1{\le}i{\le}m$, the service rate in every queue of the Y-system is smaller than the service rate in the X-system, meaning that a customer of the Y-system needs more time to run through the system than a customer of the X-system. This can be expressed with the following result: for $k \le p$, for all $t \ge 0$,

$$S_{Y,k}^N(t) \overset{\text{def.}}{=} \sum_{i=1}^{k} Y_i^N(t) \ge \sum_{i=1}^{k} X_{2i}^N(t) \overset{\text{def.}}{=} S_{X,k}^N(t). \tag{3.12}$$

This can be proven by induction, using simple coupling arguments. The cases $k = 1$ is straightforward. Assume that we have the result for $k-1$, the only Poisson processes that change $(S_k^N(t))$ are $\mathcal{P}_1$ and $\mathcal{P}_{2k+1}$. Note that $\mathcal{P}_1$ is more likely for $(S_{Y,k}(t))$ than $(S_{X,k}(t))$, so that it preserves Relation (3.12).

For $t{>}0$, if Relation (3.12) holds on $[0,t)$ and a jump of $\mathcal{P}_{2k+1}$ occurs at time t, there are two possible situations:

- If $S_{Y,k}(t{-}) \ge S_{X,k}(t{-}) + 1$, the relation will still be valid after one jump.
- If $S_{Y,k}(t{-}) = S_{X,k}(t{-})$, since $S_{Y,k-1}(t{-}){\ge}S_{X,k-1}(t{-})$ by hypothesis of induction we have $Y_k(t{-}){\le}X_{2k}(t{-})$, along with the fact that $X_{2k+1}(t{-}){\le}\eta N$, we know that

$$\mathcal{P}_{2k+1}((0,\kappa_{2k+1}X_{2k}^N(t{-})X_{2k+1}^N(t{-})),\mathrm{d}t) \ge \mathcal{P}_{2k+1}((0,\kappa_{2k+1}\eta N Y_k^N(t{-})),\mathrm{d}t),$$

which proves that Relation (3.12) is still true after the jump.
This concludes the proof of Proposition 3.4. □

THEOREM 3.5. *If $m{=}2p{+}1$ and $\alpha{\in}H_0$ such that $\alpha_{2i}{=}0$ and $\alpha_{2i-1}{>}0$, for all $1{\le}i{\le}p{+}1$ then, under Condition* (3.5)*, for the convergence in distribution, the relations*

$$\lim_{N\to+\infty}\left(\frac{X^N_{2i}(t)}{N},1{\le}i{\le}p\right){=}(0),\ \lim_{N\to+\infty}\left(\frac{X^N_{2i+1}(t)}{N},0{\le}i{\le}p\right)=(\ell(t))=(\ell_{2i+1}(t))$$

hold, where $(\ell(t))$ is the solution of the system of ODEs, for $1{\le}i{\le}p{-}1$,

$$\text{(3.13)}\qquad \begin{cases}\dot\ell_1(t) &= -\kappa_1\dfrac{\kappa_2}{\kappa_3}\dfrac{\ell_1(t)}{\ell_3(t)}\ell_1(t),\\ \dot\ell_{2i+1}(t) &= \kappa_1\left(\dfrac{\kappa_{2i}}{\kappa_{2i+1}}\dfrac{\ell_{2i-1}(t)}{\ell_{2i+1}(t)}-\dfrac{\kappa_{2(i+1)}}{\kappa_{2i+3}}\dfrac{\ell_{2i+1}(t)}{\ell_{2i+3}(t)}\right)\ell_1(t),\\ \dot\ell_{2p+1}(t) &= \kappa_1\dfrac{\kappa_{2p}}{\kappa_{2p+1}}\dfrac{\ell_{2p-1}(t)}{\ell_{2p+1}(t)}\ell_1(t){-}\kappa_{2p+2}\ell_{2p+1}(t),\end{cases}$$

with initial point (α_{2i+1}). The function $(\ell(t))$ converges to (0) as t goes to infinity.

Hence from the large state $(\alpha_1,0,\alpha_3,0,\dots,0,\alpha_m)N$, the process $(X^N(t))$ returns on the normal timescale to a neighborhood of the origin along the curve $(\ell_1(t),0,\ell_3(t),0,\dots,0,\ell_m(t))N$ of $\mathbb{R}^m_+$.

PROOF. Since we are interested in the order of magnitude in N of the components of the vector $(X^N_{2i+1}(t),0{\le}i{\le}p)$ on a finite time interval, the external arrivals with rate κ_0 do not play any role. Therefore, without a loss of generality, we can assume that $\kappa_0{=}0$, in particular $(X^N_1(t){+}X^N_3(t){+}\cdots{+}X^N_{2p+1}(t))$ is a non-increasing process by Relations (3.6). The assumptions of Proposition 3.4 are therefore satisfied, the notations for T_N, η and λ in this proposition and its proof are used.

The proof is achieved in three steps. The main difficulty is controlling that the coordinates with an even index are $O(1)$ while the others are of the order of N.

STEP 1. TIGHTNESS OF THE OCCUPATION MEASURE.
Let the *stopped occupation measure* μ_N on $\mathbb{R}_+{\times}\mathbb{N}^p$ is defined by

$$\langle\mu_N,f\rangle=\int_0^{T_N}f(s,(X^N_{2i}(s),1{\le}i{\le}p))\,\mathrm{d}s,$$

for any $f{\in}\mathcal{C}_c(\mathbb{R}_+{\times}\mathbb{N}^p)$, where T_N is defined by Relation (3.11). We now show that(μ_N) is tight for the convergence in distribution.

The tightness of the sequence of random measures (μ_N) follows from Proposition 3.4, Relation (3.9) of Proposition 3.3 and Lemma 1.3 of Kurtz [55]. Lemma 1.4 of this reference gives that any limiting point of μ_∞ of a subsequence (μ_{N_k}) can be expressed as

$$\text{(3.14)}\qquad \langle\mu_\infty,f\rangle=\int_0^{+\infty}\int_{\mathbb{N}^p}f(s,x)\pi_s(\mathrm{d}x)\,\mathrm{d}s,$$

for $f{\in}\mathcal{C}_c(\mathbb{R}_+{\times}\mathbb{N}^p)$, where (π_s) is an optional measure-valued process with values in the subset of probability distributions on $\mathbb{R}_+{\times}\mathbb{N}^p$.

If f is a continuous function on $\mathbb{R}_+{\times}\mathbb{N}^p$ such that $|f(t,x)|{\le}C(t)\|x\|$ for a continuous function $(C(t))$, $x{=}(x_i){\in}\mathbb{N}^p$ and $\|x\|{=}x_1{+}\cdots{+}x_p$, then, for the convergence in distribution of processes, the relation

$$\text{(3.15)}\qquad \lim_{k\to+\infty}\left(\int_0^{T_{N_k}}f(s,(X^{N_k}_{2i}(s)))\,\mathrm{d}s\right)=\left(\int_0^{+\infty}\int_{\mathbb{N}^p}f(s,x)\pi_s(\mathrm{d}x)\,\mathrm{d}s\right).$$

holds. The only (small) difficulty to show this result is the local behavior of $(X_N(t))$ at $t{=}0$. For $\varepsilon{>}0$, we can show that

$$\lim_{k\to+\infty}\left(\int_\varepsilon^{T_{N_k}} f(s,(X_{2i}^{N_k}(s)))\,\mathrm{d}s\right)=\left(\int_\varepsilon^{+\infty}\int_{\mathbb{N}^p} f(s,x)\pi_s(\mathrm{d}x)\,\mathrm{d}s\right). \tag{3.16}$$

The proof is standard, by using the convergence of (μ_{N_k}), the criterion of the modulus of continuity, Theorem 7.3 of Billingsley [14], and an equivalent result to Relation (3.9), which is easily proved:

$$\sup_{N\geq 1}\sup_{t>\varepsilon}\mathbb{E}\left(\left(\sum_{i=1}^p Y_i^N(Nt)\right)^2\right)<+\infty. \tag{3.17}$$

Now we have to show that for η, $\eta'{>}0$, we can find $\varepsilon{>}0$ and $k_0{\geq}1$ such that for all $k{\geq}k_0$,

$$\mathbb{P}\left(\left|\int_0^\varepsilon f(s,(X_{2i}^{N_k}(s)))\,\mathrm{d}s\right|>\eta\right)<\eta'. \tag{3.18}$$

We use the same notations as in the proof of Proposition 3.3, and

$$S_p^0\overset{\text{def.}}{=}E_{1,1}^0+\cdots+E_{p,1}^0.$$

For $\varepsilon{\in}(0,1)$ and $\delta{>}0$, we have

$$\begin{aligned}\mathbb{E}\left(\left|\int_0^\varepsilon f(s,(X_{2i}^{N_k}(s)))\,\mathrm{d}s\right|\right)&\leq\overline{C}\int_0^\varepsilon\mathbb{E}(\|(X_{2i}(s))_i\|)\,\mathrm{d}s\\&\leq\overline{C}\lambda\mathbb{E}(S_p)\varepsilon+\overline{C}\int_0^\varepsilon\|(X_{2i}(0))_i\|\mathbb{P}\left(S_p^0\geq sN\right)\mathrm{d}s\\&\leq\overline{C}\lambda\mathbb{E}(S_p)\varepsilon+\overline{C}\frac{\|(X_{2i}(0))_i\|}{\delta N}\mathbb{E}\left(e^{\delta S_p^0}\right)(1{-}e^{-\varepsilon N\delta}),\end{aligned}$$

with $\overline{C}{=}\max\{C(t),t{\in}[0,1]\}$, which leads to Equation (3.18). By letting ε go to 0, we obtain Relation (3.15).

Step 2. A lower bound for T_N.

For $1{\leq}i{<}p$, let

$$(D_i^N(t))\overset{\text{def.}}{=}\left(\int_{(0,t]}\mathcal{P}_{2i}\left(\left(0,\kappa_{2i}X_{2i-1}^N(s{-})X_{2i}^N(s{-})\right),\mathrm{d}s\right)\right),$$

then $D_i(t)$ is the number of molecules which have been transformed from S_{2i-1} into S_{2i+1} up to time t. When $t{<}T_N$, Proposition 3.4 and its proof give the relation

$$D_i^N(t)\leq D_{Y,i}^N(t)\overset{\text{def.}}{=}\int_{(0,t]}\mathcal{P}_{2i}\left(\left(0,\kappa_{2i}\lambda N\|Y^N(Ns{-})\|\right),\mathrm{d}s\right),$$

Again standard arguments of stochastic calculus and the ergodic theorem for the Markov process $(Y(t))$ give the convergence in distribution

$$\lim_{N\to+\infty}\left(\frac{D_{Y,i}^N(t)}{N}\right)=\left(\lambda^2\kappa_{2i}t\sum_{j=1}^p\frac{1}{\mu_j}\right).$$

If t_0 is chosen so that for all $1{\leq}i{<}p$,

$$t_0{<}\frac{\alpha_{2i-1}}{2\lambda^2\kappa_{2i}\sum_{j=1}^p 1/\mu_j},$$

and $t_0{<}\alpha_{2p+1}/(2\lambda\kappa_{2p+2})$, then the sequence $(\mathbb{P}(T_N{\geq}t_0))$ converges to 1.

The SDEs for $(X_N(t))$ give the relation

$$\overline{X}^N_{2i+1}(t) = \frac{x^N_{2i+1}}{N} + M^N_i(t) + \kappa_{2i}\int_0^t \overline{X}^N_{2i-1}(s)X^N_{2i}(s)\,\mathrm{d}s - \kappa_{2(i+1)}\int_0^t \overline{X}^N_{2i+1}(s)X^N_{2(i+1)}(s)\,\mathrm{d}s, \tag{3.19}$$

where $(M^N_i(t))$ is a martingale whose previsible increasing process is

$$\left(\left\langle M^N_i(t)\right\rangle\right) = \left(\frac{1}{N}\left(\kappa_{2i}\int_0^t \overline{X}^N_{2i-1}(s)X^N_{2i}(s)\,\mathrm{d}s + \kappa_{2(i+1)}\int_0^t \overline{X}^N_{2i+1}(s)X^N_{2(i+1)}(s)\,\mathrm{d}s\right)\right).$$

On the time interval $[0,t_0]$, with Doob's Inequality, since $\overline{X}^N_{2i\pm1}(s)\leq\lambda$ for $s\leq t_0$, the convergence in distribution (3.15) gives that $(M^{N_k}_i(t))$ is converging in distribution to (0). The criterion of the modulus of continuity, Theorem 7.3 of Billingsley [14], with Relation (3.19) shows that the sequence of processes $(\overline{X}^o_N(t)) \overset{\text{def.}}{=} (\overline{X}^N_{2i+1}(t), 0\leq i\leq p)$ is tight for the convergence in distribution on $[0,t_0]$. Without loss of generality one can assume that the subsequence (N_k) is such that $(\mu_{N_k}, (\overline{X}^o_{N_k}(t)))$ is converging to $(\mu_\infty, (\ell(t)))$, where μ_∞ has the representation (3.14).

Step 3. Identification of μ_∞.
In this part the convergence in distribution of processes refers implicitly to the time interval $[0,t_0]$. Let g a function on $\mathbb{N}^p$ with finite support on $\mathbb{N}^p$. Relations (3.6) gives the relation, if $(X^e_N(t)) \overset{\text{def.}}{=} (X^N_{2i}(t), 1\leq i\leq p)$,

$$g\left(X^e_N(t)\right) = g\left(X^e_N(0)\right) + M^N_g(t) + \kappa_1\int_0^t N\nabla_{e_2}(g)\left(X^e_N(s)\right)\overline{X}^N_1(s)\,\mathrm{d}s + \sum_{i=1}^p \kappa_{2i+1}\int_0^t N\nabla_{(e_{2(i+1)}-e_{2i})}(g)\left(X^e_N(s)\right)\overline{X}^N_{2i+1}(s)X^N_{2i}(s)\,\mathrm{d}s, \tag{3.20}$$

where, for $1\leq k\leq m=2p+1$, e_k is the kth unit vector of $\mathbb{R}^m_+$, with the convention $e_{2(p+1)}=0$, and, for $a\in\mathbb{R}^p_+$,

$$\nabla_a(g)(x) = g(x+a)-g(x), \quad x\in\mathbb{R}^p_+.$$

With the same argument as for the martingale $(M^N_i(t))$ above, it is easily seen that $(M^{N_k}_g(t)/N)$ is converging in distribution to the null process. By dividing by N Relation (3.20), we obtain the relation

$$\left(\int_0^t\int_{N^p} \nabla_{e_2}(f)\,(u)\,\kappa_1\ell^N_1(s)\pi_s(\mathrm{d}u)\,\mathrm{d}s + \int_0^t\int_{N^p}\sum_{i=1}^p \nabla_{(e_{2(i+1)}-e_{2i})}(f)\,(u)\,\kappa_{2i+1}\ell_{2i+1}(s)u_i\pi_s(\mathrm{d}u)\,\mathrm{d}s\right) = (0),$$

which can be written as

$$\left(\int_0^t\int_{N^p} R_{\lambda(s),\mu(s)}(f)(u)\pi_s(\mathrm{d}u)\,\mathrm{d}s\right) = (0),$$

with, for $s\geq0$, $\lambda(s)=\kappa_1\ell^N_1(s)$ and $\mu(s)=(\mu_i(s))$ and $\mu_i(s)=\kappa_{2i+1}\ell_{2i+1}(s)$ for $1\leq i\leq p$. Step 2 gives that, almost surely for all $s\leq t_0$, the relation $\ell_{2i+1}(s)\geq\eta$ holds. From

there, Proposition 3.3 and with standard arguments, as Lemma 1.5 of Kurtz [55], see also Section B.2.2 of Laurence and Robert [52], we obtain that, for $h\in\mathcal{C}_c([0,t_0]\times\mathbb{N}^p)$,

$$\int_{\mathbb{R}_+\times\mathbb{N}^p} h(s,u)\mu_\infty(\mathrm{d}s,\mathrm{d}u)=\int_0^{+\infty}\int_{\mathbb{N}^p} h(s,u)\prod_{i=1}^{p}\mathrm{Pois}\left(\frac{\kappa_1\ell_1^N(s)}{\kappa_{2i+1}\ell_{2i+1}(s)}\right)(\mathrm{d}u_i)\,\mathrm{d}s,$$

where $\mathrm{Pois}(\rho)$ is the Poisson distribution with parameter $\rho>0$.

Relation (3.19) on the subsequence (N_k) gives, as k go to infinity, the relation

$$\begin{aligned}
\ell_1(t)&=\alpha_1-\int_0^t\kappa_2\ell_1(s)\frac{\kappa_1\ell_1(s)}{\kappa_3\ell_3(s)}\,\mathrm{d}s,\\
\ell_{2i+1}(t)&=\alpha_{2i+1}+\int_0^t\kappa_{2i}\ell_{2i-1}(s)\frac{\kappa_1\ell_1(s)}{\kappa_{2i+1}\ell_{2i+1}(s)}\,\mathrm{d}s\\
&\qquad-\int_0^t\kappa_{2(i+1)}\ell_{2i+1}(s)\frac{\kappa_1\ell_1(s)}{\kappa_{2i+3}\ell_{2i+3}(s)}\,\mathrm{d}s,\quad 1\le i\le p-1,\\
\ell_{2p+1}(t)&=\alpha_{2p+1}+\int_0^t\kappa_{2p}\ell_{2p-1}(s)\frac{\kappa_1\ell_1(s)}{\kappa_{2p+1}\ell_{2p+1}(s)}\,\mathrm{d}s\\
&\qquad-\int_0^t\kappa_{2p+2}\ell_{2p+1}(s)\,\mathrm{d}s,
\end{aligned}$$

holds almost surely for all $t\in[0,t_0]$.

The convergence in distribution of $(\overline{X}_N^e(t))$ to 0 on $[0,t_0]$ is a direct consequence of b) of Proposition 3.3, Proposition 3.4 and Step 2.

Due to local Lipschitz properties, the solution of the differential system (3.13) can be defined on a maximal time interval $[0,T_\infty)$, for some $T_\infty\in\mathbb{R}_+\cup\{+\infty\}$. Assume that $T_\infty<+\infty$. For any $k=2i+1\le m$, $0\le i\le p$, by summing-up the $i+1$ first ODEs, we obtain that the function $(\ell_1(t)+\ell_2(t)+\cdots+\ell_k(t))$ is non-increasing on $[0,T_\infty)$, and, consequently, that $(\ell_k(t))$ has a limit, denoted as $\ell(T_\infty)$, when $t\nearrow T_\infty$. The last ODE of Relations (3.13), we have $\dot{\ell}_m(t)\ge-\kappa_{m+1}\ell_m(t)$, since $\alpha_m=\ell_m(0)>0$, this implies that $\ell_m(T_\infty)>0$. The same argument applied to the ODE for $(\ell_{2p-1}(t))$ gives that $\ell_{2p-1}(T_0)>0$. By induction we obtain that the relation $\ell_{2i+1}(T_\infty)>0$ holds for all $0\le i\le p$. This is a contradiction with that maximal property of T_∞. We conclude that $T_\infty=+\infty$ and, therefore, that the convergence in distribution we have obtained holds in fact on $\mathbb{R}_+$. The fact that $\ell(\infty)=(0)$ is proved with analogous arguments. The theorem is proved. □

4. The Three Species CRN

In this section, we discuss in a simple setting the process associated to the CRN, starting from large initial states from a scaling perspective. As it will be seen, there are several scaling behaviors depending on the initial state. Proposition 3.6 shows that a first order limit associated to a class of initial states is a continuous but random function. This is due also to a boundary behavior of the law of mass action.

$$\emptyset\xrightarrow{\kappa_0}S_1\xrightarrow{\kappa_1}S_1+S_2\xrightarrow{\kappa_2}S_2+S_3\xrightarrow{\kappa_3}S_3\xrightarrow{\kappa_4}\emptyset$$

The set of SDEs for this network is

$$\begin{cases}
\mathrm{d}X_1^N(t)=\mathcal{P}_0\left((0,\kappa_0),\mathrm{d}t\right)-\mathcal{P}_2\left(\left(0,\kappa_2X_1^N(t-)X_2^N(t-)\right),\mathrm{d}t\right),\\
\mathrm{d}X_2^N(t)=\mathcal{P}_1\left(\left(0,\kappa_1X_1^N(t-)\right),\mathrm{d}t\right)-\mathcal{P}_3\left(\left(0,\kappa_3X_2^N(t-)X_3^N(t-)\right),\mathrm{d}t\right),\\
\mathrm{d}X_3^N(t)=\mathcal{P}_2((0,\kappa_2X_1^NX_2^N(t-)),\mathrm{d}t)-\mathcal{P}_4((0,\kappa_4X_3^N(t-)),\mathrm{d}t),
\end{cases}\tag{3.21}$$

With the convention that N, resp. $\emptyset$, stands for a quantity of the order of N, resp. $o(N)$, Proposition 3.2 shows that the interesting initial states are $\mathcal{I}_1{=}(N,\emptyset,N)$, $\mathcal{I}_2{=}(\emptyset,\emptyset,N)$, $\mathcal{I}_3{=}(\emptyset,N,\emptyset)$, and $\mathcal{I}_4{=}(N,\emptyset,\emptyset)$. We will denote $\mathcal{I}_0{=}(\emptyset,\emptyset,\emptyset)$.

The subset $\mathcal{I}_1$ has been taken care of in Theorem 3.5. For an initial state of $\mathcal{I}_2$, with $X_3(0){\sim}x_3N$, on the time interval $[0,T]$, the first two coordinates remain $o(N)$ and it is not difficult to see that the process $(X_3^N(t)/N)$ converges in distribution to $x_3\exp(-\kappa_4 t)$.

We now discuss the remaining cases, for simplicity we focus on two initial states $(0,N,0)$ in $\mathcal{I}_3$ and $(N,0,0)$ in $\mathcal{I}_4$.

4.1. Initial State $(\mathbf{0},\mathbf{N},\mathbf{0})$. The next result shows that the limit of $(X_2^N(t)/N)$ is a continuous function, but it is random, driven by an $M/M/\infty$ process. Note that since $(L(t))$ is converging in distribution to a Poisson random variable with parameter κ_0/κ_3, the ergodic theorem gives that $(\ell_2(t))$ is almost surely converging to 0.

PROPOSITION 3.6. *If $(X_N(t))$ is the solution of (3.21) with initial condition $(0,N,0)$, then the convergence in distribution*

$$\lim_{N\to+\infty}\left(\frac{X_1^N(t)}{N},\frac{X_2^N(t)}{N},\frac{X_3^N(t)}{N}\right)=(0,\ell_2(t)),0),$$

holds, with

$$(\ell_2(t))\stackrel{\text{def.}}{=}\left(\exp\left(-\kappa_3\int_0^t L(s)\,\mathrm{d}s\right)\right),$$

where $(L(t))$ has the same distribution as the jump process associated to an $M/M/\infty$ queue with input rate κ_0 and service rate κ_4.

PROOF. We give a sketch of the proof. Some of the arguments used are the same as in the proof of Theorem 3.5. We fix some $T{>}0$ and $\delta{\in}(0,1)$, and set

$$T_2^N\stackrel{\text{def.}}{=}\inf\left\{t{>}0:X_2^N(t)\le\delta N\right\},$$

Define $(Y(t)){=}(Y_1(t),Y_2(t))$ the Markov process on $\mathbb{N}^2$ of Proposition 3.3, whose Q-matrix is $R_{\kappa_0,(\kappa_2,\kappa_4)}$. If N_0 is sufficiently large so that $\delta N_0{>}1$, then for $N{\ge}N_0$, in a similar way as in the proof of Proposition 3.4, there exists a coupling of $(X_N(t))$ and $(Y(t))$ such that, for $t{<}T_2^N$, the relation

$$X_1^N(t){+}X_3^N(t)\le Y_1(t){+}Y_2(t).$$

The SDEs (3.21) give that, for $t{\ge}0$, the relations

$$X_1^N(t)\le\mathcal{P}_0((0,\kappa_0),(0,t]) \tag{3.22}$$

$$\begin{aligned}\frac{X_2^N(t)}{N}&\le 1{+}\frac{M_1^N(t)}{N}+\frac{\kappa_1}{N}\int_0^t\mathcal{P}_0((0,\kappa_0){\times}(0,s))\,\mathrm{d}s,\\ \frac{X_2^N(t)}{N}&\ge 1+\frac{M_2^N(t)}{N}-\kappa_3\int_0^t\frac{X_2^N(s)}{N}X_3^N(s)\,\mathrm{d}s,\end{aligned} \tag{3.23}$$

hold, where $(M_i^N(t)/N)$, $i{=}1,2$, is a martingale which is converging in distribution to 0. For any ε and $\eta{>}0$, the exists N_0 such that if $N{\ge}N_0$, then the relation

$$\begin{aligned}&\frac{X_2^N(t{\wedge}T_2^N)}{N}\ge 1{-}\eta-\kappa_3\int_0^{t\wedge T_2^N}\frac{X_2^N(s)}{N}X_3^N(s)\,\mathrm{d}s\\ &\ge 1{-}\eta{-}\kappa_3\int_0^{t\wedge T_2^N}(1{+}\eta)(Y_1(s){+}Y_2(s))\,\mathrm{d}s\ge 1{-}\eta{-}\kappa_3\int_0^t(1{+}\eta)(Y_1(s){+}Y_2(s))\,\mathrm{d}s,\end{aligned}$$

holds with probability at least $1-\varepsilon$. Let

$$\tau = \sup\left\{t>0 : \int_0^t (Y_1(s)+Y_2(s))\,\mathrm{d}s \le \frac{1-\delta-\eta}{\kappa_3(1+\eta)}\right\}$$

then

$$\lim_{N\to+\infty} \mathbb{P}\left(\tau \le T_2^N\right) = 1.$$

With the criterion of the modulus of continuity, it is easily seen that the sequence $(X_2^N(t\wedge\tau\wedge T)/N)$ is tight for the convergence in distribution. We can take a subsequence associated to (N_k) converging to some continuous process $(x_2(t))$.

With the same notations as before,

$$\frac{X_2^N(t)}{N} = 1 + \frac{M_3^N(t)}{N} + \kappa_1 \int_0^t \frac{X_1^N(s)}{N}\,\mathrm{d}s - \kappa_3 \int_0^t \frac{X_2^N(s)}{N} X_3^N(s)\,\mathrm{d}s, \tag{3.24}$$

where $(M_3^N(t)/N)$ is a martingale which is converging in distribution to 0. Relation (3.22) gives that, for the convergence in distribution,

$$\lim_{N\to+\infty} \left(\int_0^t \frac{X_1^N(s)}{N}\,\mathrm{d}s\right) = (0).$$

We define $(L(t))$ as the solution of the SDE

$$\mathrm{d}L(t) = \mathcal{P}_0\left((0,\kappa_0),\mathrm{d}t\right) - \mathcal{P}_4((0,\kappa_4 L(t-)),\mathrm{d}t),$$

Now using that, on the time interval $[0,\tau\wedge T]$, the process $(X_2^N(t)/N)$ is greater than δ, so that a species S_1 is transformed into a species S_3 after a duration of time stochastically bounded by an exponential distribution with parameter $\kappa_2\delta N$. Hence, the jumps $+1$ of the process $(X_3^N(t))$ are "almost" Poisson with rate κ_0, so that $(X_3^N(t))$ is "close" to $(L(t))$. This can be stated rigorously as follows. There are at most $\mathcal{P}_0((0,\kappa_0)\times(0,T))$ jumps of size 1 for $(X_3^N(t))$, using Relation (3.23), one gets that the variable

$$\int_0^{T_2^N\wedge T} \frac{X_2^N(s)}{N}\left|X_3^N(s) - L(s)\right|\mathrm{d}s$$

converges in distribution to 0 since the duration between jumps of size $+1$ or -1 of both processes is converging to 0. By taking the limit in Relation (3.24) along the sequence (N_k) we get the identity

$$(x_2(t), t\le\tau\wedge T) = \left(\int_0^t x_2(s)L(s)\,\mathrm{d}s, t\le\tau\wedge T\right),$$

which gives the desired convergence of $(X_2^N(t)/N)$ on the time interval $[0,\tau\wedge T]$ and the representation of its limit. The procedure is repeated starting at time $\tau\wedge T$. The proposition is proved. □

4.2. Initial State $(\mathbf{N},\mathbf{0},\mathbf{0})$. When the initial state is $(N,0,0)$, guessing the time evolution is a little more tricky. One can see that $(X_2^N(t))$ grows quickly, at rate $\kappa_1 N$ initially, and then $(X_3^N(t))$ grows at rate $\kappa_2 X_1^N(t)X_2^N(t)$. The problem is of understanding the correct orders of magnitude of $(X_2^N(t), X_3^N(t))$ since all these reactions interact. A (vague) intuition suggests that, quite quickly, $(X_3^N(t))$ is of the order of N and that $(X_2^N(t))$ is $o(N)$, so that we are in the case $(N,\emptyset,N)$ already studied.

To show that this intuitive picture is "correct", we can try to use a convenient scaling such as in the following proposition.

PROPOSITION 3.7. *If the initial state is* $(N,0,0)$*, then, for the convergence in distribution,*

$$\lim_{N\to+\infty}\left(\frac{X_1^N(t/N^{2/3})}{N},\frac{X_2^N(t/N^{2/3})}{N^{1/3}},\frac{X_3^N(t/N^{2/3})}{N^{2/3}},t\geq 0\right)=(1,x_2(t),x_3(t),t\geq 0)$$

where $(x_2(t),x_3(t))$ *is solution of the ODE*

$$\begin{cases}\dot{x}_2(t)=\kappa_1-\kappa_3x_2(t)x_3(t)\\ \dot{x}_3(t)=\kappa_2x_2(t).\end{cases}$$

We skip the proof of this result since the arguments are standard. The solution of the above ODE is such that, as t goes to infinity, $x_2(t)$, resp. $x_3(t)$ converges to 0, resp. $+\infty$, which confirm our intuition. Note that the first coordinate of the process does not change at all. This is nevertheless not enough to obtain rigorously the correct orders of magnitude, for the $(X_3^N(t))$ in particular. This shows that a scaling analysis, in the sense of deriving the convergence in distribution of scaled sample paths, is not always the best approach.

Instead, with a combination of coupling arguments and convergence results, in a formulation similar as the one used in Proposition 3.2, we can prove that at some instant, deterministic here, the coordinates of the process have the right order of magnitude.

The following proposition shows that if the initial state is $(N,0,0)$, then, at any time $t_0>0$, the state of the CRN satisfies the conditions of the initial state of Theorem 3.5. It is in fact immediately in $\mathcal{I}_1$.

PROPOSITION 3.8. *If* $(X_N(t))$ *is the solution of* (3.21) *with initial condition* $(N,0,0)$*, then for any* $\varepsilon>0$ *and* $t_0>0$*, there exists* $K>0$ *and* $\eta>0$ *such that*

$$\liminf_{N\to+\infty}\mathbb{P}\left(\frac{X_1^N(t_0)}{N},\frac{X_3^N(t_0)}{N}\in(\eta,1],X_2^N(t_0)\leq K\right)\geq 1-\varepsilon.$$

PROOF. Since we are interested in the order of magnitude of $(X_1^N(t))$ in N on finite time intervals, as before, without loss of generality, we can assume that $\kappa_0=0$. In the rest of the proof, the first term of the right-hand side of the first relation of the SDEs (3.21) is removed. It is then easily seen that, in this case, the process $(X_1^N(t)+X_3^N(t))$ is non-increasing, in particular $X_3^N(t)\leq N$ for all $t\geq 0$.

For $\delta\in(0,1)$, we introduce the stochastic process $(Y_1^N(t),Y_2^N(t),Y_3^N(t))$, as the solution of the SDE

$$\begin{cases}\mathrm{d}Y_1^N(t) &= -\mathcal{P}_2((0,\kappa_2Y_1^NY_2^N(t-)),\mathrm{d}t)\\ \mathrm{d}Y_2^N(t) &= \mathcal{P}_1((0,\kappa_1\delta N),\mathrm{d}t)-\mathcal{P}_3((0,\kappa_3NY_2^N(t-)),\mathrm{d}t)\\ \mathrm{d}Y_3^N(t) &= \mathcal{P}_2((0,\kappa_2\delta NY_2^N(t-)),\mathrm{d}t)-\mathcal{P}_4((0,\kappa_4Y_3^N(t-)),\mathrm{d}t),\end{cases}\tag{3.25}$$

with $Y_1^N(0)=N$, $Y_2^N(0)=Y_3^N(0)=0$.

Define

$$T_1^N=\inf\{t\geq 0:X_1^N(t)\leq\delta N\}.$$

By induction on the jumps of the processes $(X_N(t),Y_N(t))$ on the time interval, it is not difficult to prove that, for any $t<T_1^N$, the relations $X_1^N(t)\geq Y_1^N(t)$, $X_2^N(t)\geq Y_2^N(t)$ and $X_3^N(t)\geq Y_3^N(t)$ hold.

The process $(Y_2^N(t))$ has the same distribution as $(L(Nt))$ where $(L(t))$ is the process of an $M/M/\infty$ queue with arrival rate $\kappa_1\delta$ and service rate κ_3. Hence, with the ergodic theorem for positive recurrent Markov processes, we get that, for the convergence in distribution,

$$\lim_{N\to+\infty}\left(\int_0^t f\left(Y_2^N(s)\right)\mathrm{d}s\right)=\left(\left\langle\mathrm{Pois}\left(\frac{\kappa_1\delta}{\kappa_3}\right),f\right\rangle t\right),$$

for any function f on $\mathbb{N}$ with finite support.

By using the classical approach to the proof of a stochastic averaging principle as presented in Kurtz [55], see also Section B.2.2 of [52] for an example, we obtain the following convergence in distribution

$$\lim_{N\to+\infty}\left(\frac{Y_1^N(t)}{N},\frac{Y_3^N(t)}{N}\right)=\left(e^{-\kappa_1\kappa_2\delta t/\kappa_3},\frac{\kappa_1\kappa_2\delta^2}{\kappa_3\kappa_4}\left(1-e^{-\kappa_4 t}\right)\right).$$

For $t_0>0$, one can choose $\delta\in(0,1)$ sufficiently small so that $\delta<\exp(-\kappa_1\kappa_2\delta t_0/\kappa_3)$, the above convergence shows that

$$\lim_{N\to+\infty}\mathbb{P}\left(T_1^N\geq t_0\right)=1. \tag{3.26}$$

This concludes the proof of the lower bounds of $X_1^N(t_0)/N$ and $X_3^N(t_0)/N$. Remains to show the upper bound of $X_2^N(t_0)$.

Define $\delta'\in(0,1)$ sufficiently small so that $\delta'<\kappa_1\kappa_2\delta^2\left(1-e^{-\kappa_4 t_0}\right)/(\kappa_3\kappa_4)$, and

$$T_3^N=\inf\{t\geq 0: X_3^N(t)\leq\delta' N\}.$$

In a similarly way, as for Relation (3.26), we have

$$\lim_{N\to+\infty}\mathbb{P}\left(T_3^N\geq t_0\right)=1. \tag{3.27}$$

We introduce $(Z_2^N(t))$ the solution of the SDE,

$$\mathrm{d}Z_2^N(t)=\mathcal{P}_1((0,\kappa_1 N),\mathrm{d}t)-\mathcal{P}_3((0,\kappa_3\delta' N Z_2^N(t-)),\mathrm{d}t),$$

with initial condition $Z_2^N(0)=0$. As before, the process $(Z_2^N(t))$ has the same distribution as $(L^+(Nt))$ where $(L^+(t))$ is the process of an $M/M/\infty$ queue with arrival rate κ_1 and service rate $\kappa_3\delta'$. For any $t_0>0$, $Z_2^N(t_0)$ converges in distribution to a Poisson distribution $\mathrm{Pois}(\rho)$ with parameter $\rho\overset{\text{def.}}{=}\kappa_1/(\delta'\kappa_3)$.

Since $X_1^N(t)\leq N$ for all $t\geq 0$ and $X_3^N(t)\geq\delta' N$ for $t\leq T_3^N$, it is easily seen that $X_2^N(t)\leq Z_2^N(t)$ for $t\leq T_3^N$. We conclude the proof of the proposition with Relation (3.27). □

5. AIMD processes: Invariant Distributions and a Limit Theorem

We introduce two classes of AIMD stochastic processes (Additive Increase Multiplicative Decrease) in Sections 5.1 and 5.2 which play an important role in the limit results of Sections 6 and 7 for the CRN with four nodes. Section 5.3 gives an averaging result where fast processes are AIMD processes, it will be used to establish the averaging principle of Section 7.

The first of these AIMD processes, $(R_1(t))$, describes the asymptotic behavior of the fourth coordinate $(X_4^N(t))$ on the timescale $(\sqrt{N}t)$ when $(X_3^N(t))$ is 1. The other one, $(R_0(t))$, is associated to the asymptotic behavior of $(X_4^N(t))$ on the timescale $(\sqrt{N}t)$. The asymptotic time evolution of $(X_2^N(\sqrt{N}t)/N, X_4^N(\sqrt{N}t)/\sqrt{N})$ can be expressed in terms of these AIMD processes.

In the averaging principle proved in Section 7, the fast processes involved are AIMD processes. For this reason, the asymptotic dynamic of the system is expressed in terms of functionals of their invariant distributions. Sections 5.1 and 5.1 give an explicit expression of the invariant distributions of $(R_1(t))$ and $(R_0(t))$.

Section 5.3 establishes an asymptotic result for the time evolution of a particle system. This is a key ingredient in the proofs of limit theorems of Section 7.

Definition 3.9. *For a and $b>0$, $\Gamma_0(a,b)$ is the distribution on $\mathbb{R}_+$ with density*

$$\frac{b}{\Gamma(a)}(bx)^{a-1}e^{-bx},\quad x\geq 0.$$

The function Γ is the classical Gamma function. See Whittaker and Watson [78]. The Laplace transform of $\Gamma_0(a,b)$ at $\xi{\geq}0$ is given by

$$\left(\frac{b}{b{+}\xi}\right)^a.$$

The fractional moment of order $1/2$ of this distribution is

$$\int_0^{+\infty}\sqrt{x}\,\Gamma_0(a,b)(\mathrm{d}x)=\frac{1}{\sqrt{b}\Gamma(a)}\int_0^{+\infty}x^{a-1/2}e^{-x}\,\mathrm{d}x=\frac{1}{\sqrt{b}}\frac{\Gamma(a{+}1/2)}{\Gamma(a)}. \tag{3.28}$$

5.1. The process $(R_1(t))$. For α and $\beta{>}0$, let $\mathcal{N}_1$ be a Poisson point process on $\mathbb{R}^2{\times}[0,1]$ with intensity measure $\alpha\,\mathrm{d}s{\times}\,\mathrm{d}t{\times}\beta a^{\beta-1}\,\mathrm{d}a$. The point process $\mathcal{N}_1$ can be represented as $\mathcal{N}_1{=}(u_n,v_n,U_n^{1/\beta})$ where (u_n,v_n) is a homogeneous Poisson point process with rate α, independent of the i.i.d. sequence (U_n) of uniformly distributed random variables on $[0,1]$.

We now define $(R_1(t))$ as the solution of the SDE

$$\mathrm{d}R_1(t)=\mathrm{d}t+\int_{a\in[0,1]}(a{-}1)R_1(t{-})\mathcal{N}_1((0,R_1(t{-})],\mathrm{d}t,\mathrm{d}a), \tag{3.29}$$

with $R_1(0){=}v{\geq}0$.

The asymptotic behavior of $(R_1(t))$ is described in the following proposition.

Proposition 3.10. *The process $(R_1(t))$ converges in distribution to the random variable $R_1(\infty)$. The distribution of $R_1(\infty)^2$ is $\Gamma_0((\beta{+}1)/2,\alpha/2)$ of Definition 3.9.*

The proof of the proposition uses the embedded Markov chain of $(R_1(t))$.

Proof. Let (R_n) be the embedded Markov chain of $(R_1(t))$, if $R_0{=}v{>}0$, then

$$R_1=U^{1/\beta}\,(v{+}\tau_v)\,,$$

where U and τ_v are independent random variables, U with a uniform distribution on $[0,1]$ and τ_v is such that

$$\int_0^{\tau_v}(v{+}s)\,\mathrm{d}s=E_\alpha, \tag{3.30}$$

where E_α is an exponentially distributed random variable with parameter α, i.e. for $x{\geq}0$,

$$\mathbb{P}(\tau_v\geq x)=\exp\left(-\alpha\int_0^x(v{+}s)\,\mathrm{d}s\right).$$

We obtain that

$$R_1^2=U^{2/\beta}\left(v^2{+}2E_\alpha\right).$$

Now if (U_i) and $(E_{\alpha,i})$ are independent i.i.d. sequences with the respective distribution of U and E_α,

$$\begin{aligned}W_1&\overset{\text{def.}}{=}\sum_1^{+\infty}2E_{\alpha,i}\prod_1^i U_k^{2/\beta}\\&=U_1^{2/\beta}\left(\sum_{i=2}^{+\infty}2E_{\alpha,i}\prod_{k=2}^i U_k^{2/\beta}{+}2E_{\alpha,1}\right)=U_1^{2/\beta}\,(W_2{+}2E_{\alpha,1})\,,\end{aligned}$$

and $W_2\overset{\text{dist.}}{=}W_1$. The distribution of $\sqrt{W_1}$ is therefore invariant for the Markov chain (R_n). The variable W_1 can be expressed as

$$W_1=\sum_1^{+\infty}2E_{\alpha,i}\exp(-2t_i/\beta),$$

where, for $i{\geq}1$, $t_i{=}-\ln(U_1){-}\ln(U_2){-}\cdots{-}\ln(U_i)$.

From Proposition 1.11 of Robert [67], the marked point process $\mathcal{M}=(t_n, E_{\alpha,n})$ is Poisson with intensity $\mathrm{d}u\otimes\alpha\exp(-\alpha v)\,\mathrm{d}v$ on $\mathbb{R}_+^2$ and, since

$$W_1 = \int_{\mathbb{R}_+^2} 2v\exp(-2u/\beta)\mathcal{M}(\mathrm{d}u,\mathrm{d}v),$$

with Proposition 1.5 of [67] for the Laplace transform of $\mathcal{M}$, we get that, for $\xi\geq 0$,

$$\begin{aligned}\mathbb{E}\left(e^{-\xi W_1}\right) &= \exp\left(-\int_{\mathbb{R}_+^2}\left(1-e^{-2\xi v\exp(-2u/\beta)}\right)\alpha\exp(-\alpha v)\,\mathrm{d}v\,\mathrm{d}u\right)\\ &= \exp\left(-\int_0^{+\infty}\frac{2\xi e^{-2u/\beta}}{\alpha+2\xi e^{-2u/\beta}}\,\mathrm{d}u\right) = \left(\frac{\alpha}{\alpha+2\xi}\right)^{\beta/2}.\end{aligned}$$

The distribution of W_1 is $\Gamma_0(\beta/2,\alpha/2)$, hence with density

$$\left(\frac{\alpha}{2}\right)^{\beta/2}\frac{x^{\beta/2-1}}{\Gamma(\beta/2)}e^{-\alpha x/2},\quad x\geq 0.$$

If ν_1 denotes the invariant distribution of $(R_1(t))$, its representation in terms of the invariant distribution of the embedded Markov chain gives the relation

$$\int_{\mathbb{R}_+} f(x)\nu_1(\mathrm{d}x) = \frac{1}{\mathbb{E}(\tau_{\sqrt{W_1}})}\mathbb{E}\left(\int_0^{\tau_{\sqrt{W_1}}} f\left(\sqrt{W_1}+s\right)\mathrm{d}s\right), \tag{3.31}$$

for any non-negative Borelian function f on $\mathbb{R}_+$

We take $f(x)=\exp(-\xi x)$ for some $\xi\geq 0$, Relation (3.30) with $v=\sqrt{W_1}$ gives the relation

$$\begin{aligned}\mathbb{E}(\tau_{\sqrt{W_1}})\int_{\mathbb{R}_+} e^{-\xi x}\nu_1(\mathrm{d}x) &= \frac{1}{\xi}\left(\mathbb{E}\left(\left(e^{-\xi\sqrt{W_1}}\right)\right)-\mathbb{E}\left(e^{-\xi(\sqrt{W_1}+\tau_{\sqrt{W_1}})}\right)\right)\\ &= \frac{1}{\xi}\left(\mathbb{E}\left(e^{-\xi\sqrt{W_1}}\right)-\mathbb{E}\left(e^{-\xi(\sqrt{W_1+2E_\alpha})}\right)\right),\end{aligned}$$

where E_α is an exponentially distributed random variable with parameter α, independent of W_1. By using Fubini's formula, we have

$$\begin{aligned}\mathbb{E}\left(\tau_{\sqrt{W_1}}\right)\int_{\mathbb{R}_+} e^{-\xi x}\nu_1(\mathrm{d}x) &= \frac{1}{\xi}\left(\mathbb{E}\left(e^{-\xi\sqrt{W_1}}\right)-\mathbb{E}\left(e^{-\xi(\sqrt{W_1+2E_\alpha})}\right)\right)\\ &= \int_0^{+\infty} e^{-\xi u}\mathbb{P}\left(\sqrt{W_1}\leq u\leq\sqrt{W_1+2E_\alpha}\right)\mathrm{d}u,\end{aligned} \tag{3.32}$$

and, therefore, the density of ν_1 at u is proportional to

$$\begin{aligned}\mathbb{P}\left(\sqrt{W_1}\leq u\leq\sqrt{W_1+2E_\alpha}\right) &= e^{-\alpha u^2/2}\mathbb{E}\left(e^{\alpha W_1/2}\mathbb{1}_{\{W_1\leq u^2\}}\right)\\ &= C_0 e^{-\alpha u^2/2}\int_0^{u^2} e^{\alpha w/2}x^{\beta/2-1}e^{-\alpha w/2}\,\mathrm{d}w = C_1 e^{-\alpha u^2/2}u^\beta,\end{aligned}$$

where C_0 and C_1 are multiplicative constants. The proposition is proved. □

5.2. The process $(R_0(t))$. For $\alpha>0$, let $\mathcal{N}_0$ be a Poisson point process on $\mathbb{R}\times\mathbb{R}_+$ with intensity measure $\alpha\,\mathrm{d}t\times\exp(-a)\,\mathrm{d}a$, it can be represented as the sequence of points (u_n, E_n), where (u_n) is a Poisson process on $\mathbb{R}_+$ with rate α and (E_n) is an i.i.d. sequence of exponential random variables with parameter 1, independent of (u_n).

For γ, $\beta>0$, we define $(R_0(t))$, the solution of the SDE

$$\mathrm{d}R_0(t) = -\frac{1}{\gamma}R_0(t)\,\mathrm{d}t + \int_{a\in\mathbb{R}_+}\left(\sqrt{R_0(t-)^2+\frac{a}{\beta}}-R_0(t-)\right)\mathcal{N}_0(\mathrm{d}t,\mathrm{d}a), \tag{3.33}$$

with $R_0(0)=v\geq 0$.

It should be noted that the integral expression in the above SDE is just a Dirac measure, since there is only one "a" when there is a jump in t.

PROPOSITION 3.11. *The process $(R_0(t))$ converges in distribution as t goes to infinity to a random variable $R_0(\infty)$, such that $R_0(\infty)^2$ has the distribution $\Gamma_0(\alpha\gamma/2,\beta)$ of Definition 3.9.*

PROOF. If $(W(t))=(R_0(t)^2)$, we obtain easily with a change of variable that the relation

$$\mathrm{d}W(t) = -\frac{2}{\gamma}W(t)\,\mathrm{d}t + \frac{1}{\beta}\int_{a\in\mathbb{R}_+} a\,\mathcal{N}_0(\mathrm{d}t,\mathrm{d}a).$$

Its unique solution is given by, for $t\geq 0$,

$$W(t) = W(0)e^{-2t/\gamma} + \frac{1}{\beta}\int_0^t ae^{-2(t-s)/\gamma}\mathcal{N}_0(\mathrm{d}s,\mathrm{d}a)$$
$$\stackrel{\text{dist.}}{=} W(0)e^{-2t/\gamma} + \frac{1}{\beta}\int_0^t ae^{-2s/\gamma}\mathcal{N}_0(\mathrm{d}s,\mathrm{d}a),$$

by reversibility and invariance of the Poisson process by translation. We obtain that $(W(t))$ converges in distribution to

$$W_\infty \stackrel{\text{def.}}{=} \frac{1}{\beta}\int_0^{+\infty} ae^{-2s/\gamma}\mathcal{N}_0(\mathrm{d}s,\mathrm{d}a).$$

For $\xi\geq 0$, by using the representation of the Laplace transform of a Poisson process, see Proposition 1.5 of Robert [67], we get that

$$\mathbb{E}\left(e^{-\xi W_\infty}\right) = \exp\left(-\alpha\int_{\mathbb{R}_+^2}\left(1-\exp\left(-\frac{\xi}{\beta}ae^{-2s/\gamma}\right)\right)e^{-a}\,\mathrm{d}a\,\mathrm{d}s\right)$$
$$= \exp\left(-\alpha\int_0^{+\infty}\frac{\xi e^{-2s/\gamma}}{\beta+\xi e^{-2s/\gamma}}\,\mathrm{d}s\right) = \left(\frac{\beta}{\beta+\xi}\right)^{\alpha\gamma/2}.$$

We conclude that the distribution of W_∞ is $\Gamma_0(\alpha\gamma/2,\beta)$. □

5.3. A Limiting Result on a Particle System. We investigate the limiting behavior of the time evolution of a Markov process on $\mathbb{N}$ which is described in terms of a particle system. A process of this type plays an important role in a time-changed version of the process $(U_4^N(t))$ in Section 7. A limiting result is established, Theorem 3.12, it plays a central role in the asymptotic analysis of the CRN with four nodes in Section 7.

The kinetics are as follows: At time t, we are given G, an exponential random variable variable with parameter σ, and (E_i) an independent i.i.d. sequence of exponential random variables with parameter λ.

- Arrivals. A new particle arrives at rate $P_N(t-)$, where $(P_N(t))$ is an adapted càdlàg process.
- Departures. If there are $x\in\mathbb{N}$ particles, at rate δx, $\delta>0$, any particle $1\leq i\leq x$ such that $E_i\leq G$ is removed.

With the definition of Section 2.1, $\mathcal{N}_\sigma$ is a Poisson marked point process on the state space $\mathbb{R}_+^2\times\mathbb{R}_+\times\mathbb{R}_+^{\mathbb{N}}$ whose intensity measure is

$$\mathrm{d}s\otimes\mathrm{d}t\otimes\sigma\exp(-\sigma a)\,\mathrm{d}a\otimes Q(\mathrm{d}b),$$

where Q is the distribution on $\mathbb{R}_+^{\mathbb{N}}$ of an i.i.d. sequence (E_i) of exponential random variables with parameter λ.

The process of the number of particles, $(K_N(t))$, is defined as the solution of the SDE,

$$\begin{aligned} \mathrm{d}K_N(t) &= \mathcal{P}_0((0,P_N(t-)),\mathrm{d}t) \\ &\quad - \int_{(a,b)\in\mathbb{R}_+\times\mathbb{R}_+^{\mathbb{N}}} S(K_N(t-),a,b)\mathcal{N}\left((0,\delta K_N(t-)],\mathrm{d}t,\mathrm{d}a,\mathrm{d}b\right), \end{aligned} \tag{3.34}$$

such that $K_N(0)=w_N\in\mathbb{N}$. For $a\geq 0$ and $b=(b_i)\in\mathbb{R}_+^{\mathbb{N}}$, we denote

$$S(n,a,b) \stackrel{\text{def.}}{=} \sum_{i=1}^{n} \mathbb{1}_{\{b_i\leq a\}}. \tag{3.35}$$

It can be shown, with the criterion of the modulus of continuity, that the process $(\overline{K}_N(t))=(K_N(t)/\sqrt{N})$ has convenient tightness properties on the timescale $(t/\sqrt{N})$. Since we are interested in the asymptotic behavior of this process on the normal timescale, we will investigate the asymptotic behavior of its occupation measure μ_N defined by

$$\langle \mu_N, f\rangle = \int_0^T F\left(s,\overline{K}_N(s)\right)\mathrm{d}s = \int_0^T F\left(s,\frac{K_N(s)}{\sqrt{N}}\right)\mathrm{d}s, \tag{3.36}$$

for $F\in\mathcal{C}_c([0,T]\times\mathbb{R}_+)$.

THEOREM 3.12. *If $(P_N(t))$ is a càdlàg adapted process on $\mathbb{N}$ such that $(P_N(t)/N)$ is a bounded process converging in distribution to a continuous positive process $(p(t))$ and $(K_N(t))$ is the solution of SDE (3.34), $K_N(0)=w_N$ and the sequence (w_N/N) is bounded, then the sequence of occupation measures (μ_N) defined by Relation (3.36) converges in distribution to the measure μ_∞ defined by*

$$\int_0^T\int_0^{+\infty} F(s,x)\,\mu_\infty(\mathrm{d}s,\mathrm{d}x) = \int_0^T\int_0^{+\infty} F\left(s,\sqrt{x}\right)\Gamma_0\left(\frac{\sigma}{2\lambda}+\frac{1}{2},\frac{\delta}{2p(s)}\right)(\mathrm{d}x)\,\mathrm{d}s,$$

for $F\in\mathcal{C}_c([0,T]\times\mathbb{R}_+)$, where Γ_0 is the distribution of Definition 3.9.

PROOF. The proof is done in several steps.

STEP 1: TIGHTNESS PROPERTIES OF (μ_N).

By integration of Relation (3.34) and taking the expected value, we obtain that

$$\delta\frac{\lambda}{\lambda+\sigma}\int_0^T \mathbb{E}\left(\overline{K}_N(s)^2\right)\mathrm{d}s \leq \frac{w_N}{N} + \int_0^T \mathbb{E}\left(\frac{P_N(s)}{N}\right)\mathrm{d}s,$$

hence

$$\sup_{N\geq 1}\int_0^T \mathbb{E}\left(\overline{K}_N(s)^2\right)\mathrm{d}s < +\infty. \tag{3.37}$$

This Relation will be used frequently in the proof of the different convergence theorems of Section 7. Relation (3.37), Lemmas 1.3 and 1.4 of Kurtz [55] show that 1) the sequence (μ_N) is tight for the convergence in distribution, and 2) if μ_∞ is the limit of some converging subsequence (μ_{N_k}), then it can be represented as

$$\langle \mu_\infty, F\rangle = \int_{[0,T]\times\mathbb{R}_+} F(s,x)\gamma_s(\mathrm{d}x)\,\mathrm{d}s,$$

for $F\in\mathcal{C}_c([0,T]\times\mathbb{R}_+)$, where (γ_s) is an optional process with values in probability distributions on $\mathbb{R}_+$.

If f is a bounded continuous function on $\mathbb{R}_+$, then for $T\geq 0$,

$$\mathbb{E}\left(\int_0^T f\left(\overline{K}_N(s)\right)\overline{K}_N(s)\mathbb{1}_{\left\{\overline{K}_N(s)\geq K\right\}}\,\mathrm{d}s\right) \leq \frac{\|f\|_\infty}{K}\int_0^T \mathbb{E}\left(\overline{K}_N(s)^2\right)\mathrm{d}s,$$

hence, for the convergence in distribution

$$\lim_{N\to+\infty}\int_0^T f\left(\overline{K}_N(s)\right)\overline{K}_N(s)\,\mathrm{d}s = \int xf(x)\mathbb{1}_{\{s\le T\}}\mu_\infty(\mathrm{d}s,\mathrm{d}x) = \int_0^T\int_0^{+\infty} xf(x)\gamma_s(\mathrm{d}x)\,\mathrm{d}s.$$

The next steps are devoted to the identification of the limit.

Step 2: Control of martingales.
If f is some continuous bounded function on $\mathbb{R}_+$, then, for $t\ge 0$, standard stochastic calculus gives the relation, for $t\in[0,T]$,

$$\begin{aligned} f(\overline{K}_N(t)) &= f(\overline{K}_N(0)) + M_f^N(t) \\ &\quad + \int_0^t \left(f\left(\overline{K}_N(s)+\frac{1}{\sqrt{N}}\right) - f\left(\overline{K}_N(s)\right)\right) P_N(s)\,\mathrm{d}s \\ &\quad + \int_0^t \left(\mathbb{E}\left(f\left(\frac{1}{\sqrt{N}}\sum_1^{K_N(s)}\mathbb{1}_{\{E_i>G\}}\right)\middle|\mathcal{F}_s\right) - f\left(\overline{K}_N(s)\right)\right)\delta K_N(s)\,\mathrm{d}s, \end{aligned} \tag{3.38}$$

where $(M_f^N(t))$ is a martingale. The expectation value of the last line is taken on the random variables G and (E_i). If f is $\mathcal{C}_1$ with compact support, its predictable increasing process $(\langle M_f\rangle(t))$ is such that

$$\langle M_f\rangle(t) \le 4\|f'\|_\infty^2\int_0^t \frac{P_N(s)}{N}\,\mathrm{d}s + 4\delta\|f\|_\infty^2\int_0^t K_N(s)\,\mathrm{d}s. \tag{3.39}$$

With Relations (3.39) and (3.37), we obtain that the sequence $(\mathbb{E}(\langle M_f^N/\sqrt{N}\rangle(T)))$ is converging to 0 and, therefore, with Doob's Inequality, that the martingale $(M_f^N(t)/\sqrt{N})$ converges in distribution to 0.

Step 3: A Technical estimate.
This step is dedicated to the estimation of the height of the negative jumps of $(K_N(t))$. Let $f\in\mathcal{C}_c(\mathbb{R}_+)$, and define $\Delta_N(f)$ as

$$\int_0^T\left|\mathbb{E}\left(f\left(\overline{K}_N(s)\frac{1}{K_N(s)}\sum_1^{K_N(s)}\mathbb{1}_{\{E_i>G\}}\right) - f\left(\overline{K}_N(s)e^{-\lambda G}\right)\middle|\mathcal{F}_s\right)\right|\overline{K}_N(s)\,\mathrm{d}s.$$

The law of large numbers gives the relation

$$\lim_{K_0\to+\infty}\sup_{x\ge 0}\left|f\left(x\frac{1}{K_0}\sum_1^{K_0}\mathbb{1}_{\{E_i>G\}}\right) - f\left(xe^{-\lambda G}\right)\right| = 0$$

which holds almost surely, and by Lebesgue's Theorem in L_1. The integrand of quantity $\Delta_N(f)$, with the factor $K_N(s)$ excluded, is expressed as the sum of two terms. One with the indicator function of the set $\{\overline{K}_N(s)\ge\eta\}=\{K_N(s)\ge\eta\sqrt{N}\}$, for some small $\eta>0$. The last estimate can then be used for this term. The other term with the set $\{\overline{K}_N(s)\le\eta\}$ is negligible since the integrand is the difference of two expressions which are arbitrarily both close to $f(0)$.

Relation (3.37) and Cauchy-Shwartz's Inequality give that, for $K_1>0$,

$$\sup_N\mathbb{P}\left(\int_0^T\overline{K}_N(s)\,\mathrm{d}s\ge K_1\right) \le \frac{T}{K_1}\sup_N\sqrt{\int_0^T\mathbb{E}\left(\overline{K}_N(s)^2\right)\mathrm{d}s}.$$

By combining these results, we obtain that the sequence $(\Delta_N(f))$ converges in distribution to 0.

STEP 4: CONCLUSION
If $f\in\mathcal{C}_c^1$, by dividing Relation (3.38) by $\sqrt{N}$, and letting N go to infinity, we obtain that, almost surely, the relation

$$\int_0^t\int_{\mathbb{R}_+} p(s)f'(x)\gamma_s(\mathrm{d}x)\,\mathrm{d}s+\int_0^t\int_{\mathbb{R}_+^2}\left(f\left(xe^{-\lambda a}\right)-f\left(x\right)\right)\delta x\sigma e^{-\sigma a}\,\mathrm{d}a\gamma_s(\mathrm{d}x)\,\mathrm{d}s=0,$$

holds for all $t\geq 0$. Indeed it clearly holds almost surely for a fixed t and, therefore for all $t\in\mathbb{Q}$, consequently for all $t\geq 0$ by continuity. We have established the relation, almost surely

$$\left(\int_0^t\left\langle\mathcal{A}_{p(s)}(f),\gamma_s\right\rangle\mathrm{d}s\right)=(0),$$

with, for $x\geq 0$, $a>0$

$$\mathcal{A}_a(f)(x)=af'(x)+\delta x\int_0^1\left(f\left(xu^{\lambda/\sigma}\right)-f\left(x\right)\right)\mathrm{d}u.$$

Note that $\mathcal{A}_a$ is the infinitesimal generator of the Markov process $(R_1(at))$, where $(R_1(t))$ is defined in Section 5.1 with $\beta=\sigma/\lambda$ and $\alpha=\delta/a$. Hence, there exists a subset S_f of $[0,T]$, negligible for the Lebesgue's measure, such that, almost surely, the relation $\left\langle\mathcal{A}_{p(s)}(f),\gamma_s\right\rangle=0$ holds, for all $s\in[0,T]\backslash S_f$. Since the set of $\mathcal{C}_1$ functions with compact support on $\mathbb{R}_+$ is separable for the uniform norm, there exists S_0 of $[0,T]$, negligible for the Lebesgue's measure such that, almost surely, for any f in a dense subset of such functions the relation $\left\langle\mathcal{A}_{p(s)}(f),\gamma_s\right\rangle=0$ for all $s\in[0,T]\backslash S_0$. Proposition 9.2 of Ethier and Kurtz [24] gives that for $s\in[0,T]\backslash S_0$, γ_s is the invariant distribution of $(R_1(t))$. The proposition is proved. □

6. The Four Species CRN

We investigate the asymptotic behavior of the CRN with four nodes starting from an initial state of the form $(0,N,0,0)$ for some large N. As explained in the introduction, we did not try a complete classification of initial states from the point of view of the asymptotic behavior of $(X_N(t))$ as we have done for $m=3$. We do believe however that this is *the* interesting class of initial states, i.e. with a really unusual asymptotic behavior.

By using Filonov's Theorem, see Theorem 6 of Laurence and Robert [52], it can be proved, with some tedious but straightforward technicalities, that the associated Markov process $(X(t))$ is positive recurrent. In particular, starting from $(0,N,0,0)$ the second coordinate $(X_2^N(t))$ will eventually decrease. The goal of this section and of Section 7 is of characterizing this decay.

Up to now, we have seen that the ordinary timescale was enough to observe the decay of the norm of the process. See Theorem 3.5 and Sections 4.1 and 4.2. For this initial state, the situation is significantly different. It turns out that the convenient timescale is $(\sqrt{N}t)$ and that on this timescale the process lives in a set of states of the form (a,y_N,b,v_N) with a, $b\in\mathbb{N}$ and y_N and v_N are respectively of the order of N and $\sqrt{N}$.

We analyze the scaling properties of this CRN on the normal timescale in this section. It is shown that, with a scaling in space, the Markov process converges in distribution to a jump process under a convenient topology on the space of càdlàg functions $\mathcal{D}([0,T],\mathbb{R}_+^4)$. The limiting process is an AIMD process whose invariant distribution has been investigated in Section 5.2. The asymptotic behavior on the timescale $(\sqrt{N}t)$ is analyzed in Section 7.

$$\emptyset\xrightarrow{\kappa_0}S_1\xrightarrow{\kappa_1}S_1+S_2\xrightarrow{\kappa_2}S_2+S_3\xrightarrow{\kappa_3}S_3+S_4\xrightarrow{\kappa_4}S_4\xrightarrow{\kappa_5}\emptyset$$

The set of SDEs for the state $(X_N(t))=(X_i^N(t))$ network is

$$\begin{cases} \mathrm{d}X_1^N(t) &= \mathcal{P}_0\left((0,\kappa_0),\mathrm{d}t\right) - \mathcal{P}_2\left(\left(0,\kappa_2 X_1^N X_2^N(t-)\right),\mathrm{d}t\right), \\ \mathrm{d}X_2^N(t) &= \mathcal{P}_1\left(\left(0,\kappa_1 X_1^N(t-)\right),\mathrm{d}t\right) - \mathcal{P}_3\left(\left(0,\kappa_3 X_2^N X_3^N(t-)\right),\mathrm{d}t\right), \\ \mathrm{d}X_3^N(t) &= \mathcal{P}_2((0,\kappa_2 X_1^N X_2^N(t-)),\mathrm{d}t) - \mathcal{P}_4\left(\left(0,\kappa_4 X_3^N X_4^N(t-)\right),\mathrm{d}t\right), \\ \mathrm{d}X_4^N(t) &= \mathcal{P}_3((0,\kappa_3 X_2^N X_3^N(t-)),\mathrm{d}t) - \mathcal{P}_5((0,\kappa_5 X_4^N(t-)),\mathrm{d}t). \end{cases} \tag{3.40}$$

6.1. Scaling Properties. The rest of this section and Section 7 are devoted to the asymptotic behavior of $(X_N(t))$ when the initial state is $X_N(0)=(0,y_N,0,v_N)$, with

$$\lim_{N\to+\infty}\left(\frac{y_N}{N},\frac{v_N}{\sqrt{N}}\right)=(y,v)\in\mathbb{R}_+^2,\quad y>0. \tag{3.41}$$

The time evolution of the process $(X_N(t))$ is investigated, in a natural way, by a representation in terms of several steps of a cycle defined in terms of the points of $\mathcal{P}_0((0,\kappa_0],\mathrm{d}t)$. We give a heuristic description of it for the moment. The cycle describes, with high probability, the time evolution of the CRN in terms of the values of $(X_1^N(t),X_3^N(t))$ with the successive states $(0,0)$, $(1,0)$, $(0,1)$ and $(0,0)$. This is not a formal definition but more an (hopefully) insightful picture of an important aspect of the kinetics of our CRN.

A Heuristic Description of a Cycle.

(a) If the initial state is $(0,y_N,0,v_N)$. Let t_1 be the first point of the point process $\mathcal{P}_0((0,\kappa_0],\mathrm{d}t)$. On the time interval $[0,t_1)$, only the last coordinate $(X_4^N(t))$ changes, via the SDE

$$\mathrm{d}A_N(t) = -\mathcal{P}_5((0,\kappa_5 A_N(t-)),\mathrm{d}t), \tag{3.42}$$

with $A_N(0)=v_N$.

At time t_1 the state of the CRN is $(1,y_N,0,X_4^N(t_1))$.

(b) If the initial state is $(1,y_N,0,v_N)$, the variable τ_N^1 is the time when the 1 at the first coordinate "moves" to the third coordinate. At this instant, on the event $\{t_1>\tau_N^1\}$, the state becomes $(0,y_N+Y_N,1,v_N-W_N)$, where

$$Y_N=\mathcal{P}_1((0,\kappa_1)\times(0,\tau_N^1)) \text{ and } W_N=\int_0^{\tau_N^1}\mathcal{P}_5((0,\kappa_5 X_4^N(s-)),\mathrm{d}s).$$

It is not difficult to see that the random variable τ_N^1 is of the order of $1/N$ and, consequently, that the sequence $(Y_N/N,W_N/\sqrt{N})$ is converging in distribution to $(0,0)$. At time τ_N^1, the process starts at a state "close" to $(0,y_N,1,v_N)$.

(c) If the initial state is $(0,y_N,1,v_N)$, the second coordinate may decrease until the time τ_N^2 when the "1" of the third coordinate becomes 0. The state is at this moment $(0,X_2^N(\tau_N^2),0,X_4^N(\tau_N^2))$ with high probability.

In this approximate description, the possible values for the first and the third coordinates either 0 or 1. This turns out to be essentially an accurate asymptotic description of the CRN on the "normal" timescale (t) investigated in this section. A similar statement for the timescale $(\sqrt{N}t)$ of Section 7 still holds but requires quite different arguments.

The distribution of the duration of the first step is exponential with parameter κ_0. The decay of $(X_4^N(t))$ occurs essentially during this step. Step (2) is a (short) transition, mentioned only to have a straight formulation of the limit results.

The third step is when $(X_2^N(t))$ decreases and $(X_4^N(t))$ builds up. As it will be seen its duration is $O(1/\sqrt{N})$ and the number of jumps of the process during this step is large, of the order of $\sqrt{N}$. This feature has a significant impact on the

statements of the scaling results for $(X_2^N(t)/N, X_4^N(t)/\sqrt{N})$ via the topologies used on the space of càdlàg functions.

With a time change, the third step can be "removed" and a convergence result holds for the usual Skorohod topology, the J_1-topology. See Proposition 3.15. Otherwise, on the full timescale, the M_1-Skorohod topology or the S-topology has to be used, see Proposition 3.16. See Whitt [77] and Jakubowski [44] for general presentations of these topologies.

In any of these cases, the limiting process is an AIMD process analyzed in Section 5.2. See Figure 2. We start with the main limiting result for the third step.

PROPOSITION 3.13. *If $X_N(0)=(0, y_N, 1, v_N)$ satisfy Relation* (3.41) *and*

$$\tau_N^2 \stackrel{\text{def.}}{=} \inf\left\{t>0 : X_3^N(t) = 0\right\},$$

then, under Condition (3.41)*, the relation*

$$\lim_{N\to+\infty}\left(\sqrt{N}\tau_N^2, \frac{X_2^N(\tau_N^2)}{N}, \frac{X_4^N(\tau_N^2)}{\sqrt{N}}\right) = (H_{y,v}, y, v+\kappa_3 y H_{y,v}), \tag{3.43}$$

holds for the convergence in distribution, where $H_{y,v}$ is a non-negative random variable whose distribution is given by

$$\mathbb{E}_y(f(H_{y,v})) = \int_0^{+\infty} f\left(\frac{1}{\kappa_3 y}\left(\sqrt{v^2+2\frac{\kappa_3}{\kappa_4}ys} - v\right)\right) e^{-s}\,\mathrm{d}s, \tag{3.44}$$

for any function $f\in\mathcal{C}_c(\mathbb{R}_+)$.

PROOF. Let $(Y_2^N(t), Y_4^N(t))$ be the solution of the SDEs

$$\begin{cases} \mathrm{d}Y_2^N(t) &= -\mathcal{P}_3\left(\left(0, \kappa_3 Y_2^N(t-)\right), \mathrm{d}t\right), \\ \mathrm{d}Y_4^N(t) &= \mathcal{P}_3((0, \kappa_3 Y_2^N(t-)), \mathrm{d}t) - \mathcal{P}_5((0, \kappa_5 Y_4^N(t-)), \mathrm{d}t), \end{cases}$$

with the initial condition $(Y_2^N(0), Y_4^N(0))=(N, v_N)$. Standard stochastic calculus as in Section 4 gives the convergence in distribution as processes for the uniform topology

$$\lim_{N\to+\infty}\left(\frac{Y_2^N(t/\sqrt{N})}{N}, \frac{Y_4^N(t/\sqrt{N})}{\sqrt{N}}\right) = (y, v+\kappa_3 y t). \tag{3.45}$$

We define

$$\tau_N^Y = \inf\left\{t>0 : \int_0^t \mathcal{P}_4\left(\left(0, \kappa_4 Y_4^N(s-)\right), \mathrm{d}s\right) \neq 0\right\}.$$

Let E_1 be an exponential random variable with parameter 1, independent of $\mathcal{P}_3$ and $\mathcal{P}_5$. It is easily seen that if H_N is the solution of the relation

$$\kappa_4 \int_0^{H_N} Y_4^N(s)\,\mathrm{d}s = E_1 \tag{3.46}$$

then the relation

$$(\tau_N^Y, Y_2^N(\tau_N^Y), Y_4^N(\tau_N^Y)) \stackrel{\text{dist.}}{=} (H_N, Y_2^N(H_N), Y_4^N(H_N))$$

holds. A change of variable in Relation (3.46) gives the identity

$$\kappa_4 \int_0^{\sqrt{N}H_N} \frac{Y_4^N(s/\sqrt{N})}{\sqrt{N}}\,\mathrm{d}s = E_1.$$

Relation (3.45) shows that the sequence $(\sqrt{N}H_N)$ is tight and also that any of its limiting points H satisfies the relation

$$\kappa_4\left(vH + \frac{\kappa_3}{2}yH^2\right) \stackrel{\text{dist.}}{=} E_1,$$

and therefore the convergence in distribution of this sequence to $H_{y,v}$. The relation

$$\left(\sqrt{N}H_N, \frac{Y_2^N(H_N)}{N}, \frac{Y_4^N(H_N)}{\sqrt{N}}\right) = \left(\sqrt{N}H_N, \frac{Y_2^N(\sqrt{N}H_N/\sqrt{N})}{N}, \frac{Y_4^N(\sqrt{N}H_N/\sqrt{N})}{\sqrt{N}}\right)$$

and the convergence (3.45) show that, when N goes to infinity these random variables converge in distribution to the right-hand side of Relation (3.43).

Until time $t_1 \wedge \tau_N^2$, it is easy to see that $(X_2^N(t), X_4^N(t)) \stackrel{\text{dist.}}{=} (Y_2^N(t), Y_4^N(t))$. The proof is concluded by noting that

$$\lim_{N\to+\infty} \mathbb{P}\left(t_1 < \tau_N^Y\right) = 0,$$

since t_1 is a exponential random variable with parameter κ_0, independent of $\mathcal{P}_3$ and $\mathcal{P}_5$. □

We complement the last proposition with a technical corollary which describes the time evolution on the time interval $[0, \tau_2^N)$ of the process $(X_2^N(t), X_4^N(t))$. It is used in the proof of the convergence for the M_1-topology of Section 6.3.

COROLLARY 3.14. *With the notations and Assumptions of Proposition 3.13, if $T>0$, for the convergence in distribution*

$$\lim_{N\to+\infty} \left(\frac{X_2^N(t/\sqrt{N})}{N}, \frac{X_4^N(t/\sqrt{N})}{\sqrt{N}}, t \le \sqrt{N}\tau_2^N\right) = ((y, v+\kappa_3 yt), t \le H_{y,v}).$$

PROOF. This is a consequence of a) the coupling of the proof of Proposition 3.13, the relation $(X_2^N(t), X_4^N(t)) = (Y_2^N(t), Y_4^N(t))$ holds for $t < \tau_2^N \wedge t_1$, and b) that Relation (3.45) can be strengthened as

$$\lim_{N\to+\infty} \left(\left(\frac{Y_2^N(t/\sqrt{N})}{N}, \frac{Y_4^N(t/\sqrt{N})}{\sqrt{N}}\right), \sqrt{N}\tau_2^N\right) = ((y, v+\kappa_3 yt), H_{y,v}).$$

□

We now return to the investigation of the asymptotic behavior of

$$\left(\frac{X_2^N(t)}{N}, \frac{X_4^N(t)}{\sqrt{N}}\right),$$

when the initial state is such that $(X_2^N(0), X_4^N(0)) = (y_N, v_N)$ and Relation 3.41 holds, and $X_1^N(0)$, $X_3^N(0) \in \{0, 1\}$. We first show that, up to a time change, there is indeed a convergence in distribution for the J_1-topology. See Proposition 3.15. Without a time change, there is a convergence in distribution but for weaker topologies, the M_1-topology and the S-topology. See the discussion in Section 6.3.

6.2. Convergence with a Random Time Change. The time change considered in this section consists in removing the instants t of step (3) of Definition 6.1, i.e. when $X_3(t) \neq 0$. We introduce, for $t \ge 0$,

$$L_0^N(t) = \int_0^t \mathbb{1}_{\{X_3^N(s)=0\}} \, \mathrm{d}s \quad \text{and} \quad \ell_0^N(t) = \inf\{s \ge 0 : L_0^N(s) > t\}.$$

PROPOSITION 3.15. *If $X_N(0) = (0, y_N, 1, v_N)$ satisfy Relation (3.41), then for the convergence in distribution for the J_1-Skorohod topology,*

$$\lim_{N\to+\infty} \left(\frac{X_2^N(\ell_0^N(t))}{N}, \frac{X_4^N(\ell_0^N(t))}{\sqrt{N}}\right) = (y, V_y(t)),$$

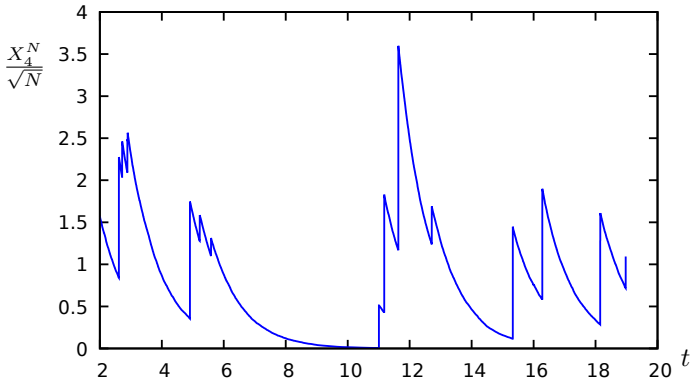


FIGURE 2. CRN with Four Nodes.
Simulation: A snapshot of $(X_4^N(t)/\sqrt{N})$.
Initial state $(0,N,0,0)$, κ_i=1, i=0,...,5 and N=9·10^6.

where $(V_y(t))$ *is the Markov process on* $\mathbb{R}_+$ *whose infinitesimal generator* $\mathcal{A}_y$ *is given by, for* $f\in\mathcal{C}_c^1(\mathbb{R}_+)$ *and* $x\in\mathbb{R}_+$,

$$\mathcal{A}_y(f)(x) = -\kappa_5 x f'(x) + \kappa_0 \int_0^{+\infty} \left(f\left(\sqrt{x^2+2\frac{\kappa_3}{\kappa_4}ys}\right) - f(x)\right) e^{-s}\,\mathrm{d}s$$

The process $(V_y(t))$ is in the class of AIMD processes introduced in Section 5.2. See Figure 2 for an illustration of a sample path of $(V_y(t))$.

PROOF. The result itself is quite intuitive in view of Proposition 3.13. Some care is nevertheless necessary in order to deal with formal aspects of the J_1-topology. The proof is direct. One of its simple ingredients is that the convergence

$$\lim_{N\to+\infty} \left(x_N \mathbb{1}_{\{a_N\le t\}} + y_N \mathbb{1}_{\{b_N\le t\}}\right) = \left(x\mathbb{1}_{\{a\le t\}} + y\mathbb{1}_{\{b\le t\}}\right) \tag{3.47}$$

holds for the J_1-topology if the sequences (x_N), (y_N),(a_N),(b_N) converge respectively to x, y, a, b with $a\neq b$. Extensions with more terms also hold. See Section VI.1 of Jacod and Shiryaev [43] for example. The proof also use Skorohod's representation theorem repeatedly. See Theorem 1.8 of Ethier and Kurtz [24].

We denote by (t_n) the non-decreasing sequence of points of $\mathcal{P}_0((0,\kappa_0],\mathrm{d}t)$, with the convention that t_0=0. An important point is that this sequence is not depending on the scaling parameter N. This property will simplify the proofs of convergence in distribution of this section. We fix T>0, as usual $\mathcal{D}([0,T])$ denotes the space of càdlàg functions on $[0,T]$, we will have to consider the events, $T\in[t_n,t_{n+1})$, $n\ge1$.

Let $(A_0^N(t))$ be the solution of SDE (3.42) with $A_0^N(0)=v_N$. A standard argument gives the convergence in distribution

$$\lim_{N\to+\infty} \left(\frac{A_0^N(t)}{\sqrt{N}}, 0\le t<T\right) = \left(ve^{-\kappa_5 t}, 0\le t<T\right), \tag{3.48}$$

for the uniform topology on $\mathcal{D}([0,T])$. By using Skorohod's representation theorem, see Theorem 1.8 of Ethier and Kurtz [24], one can assume that there exists a probability space on which the sequence of processes $(A_0^N(t))$ are defined and the convergence (3.48) holds almost surely for the uniform norm.

We use similar notations as in Definition 6.1

$$\tau_N^1 \overset{\text{def.}}{=} \inf\left\{t>0 : X_3^N(t_1+t)=1\right\}, \tau_N^2 \overset{\text{def.}}{=} \inf\left\{t>0 : X_3^N(t_1+\tau_N^1+t)=0\right\}.$$

On the event $\{t_1 \leq T\}$, the processes $((y_N, A_0^N(t)), t<t_1)$ has the same distribution as $((X_2^N(t), X_4^N(t)), t<t_1)$ and, almost surely

$$\lim_{N\to+\infty} \frac{A_0^N(t_1)}{\sqrt{N}} = \overline{v} \stackrel{\text{def.}}{=} ve^{-\kappa_5 t_1}.$$

then, by Proposition 3.13, the convergence in distribution

$$\lim_{N\to+\infty} \left(\sqrt{N}\tau_N^1, \sqrt{N}\tau_N^2, \frac{X_2^N(t_1+\tau_N^1+\tau_N^2)}{N}, \frac{X_4^N(t_1+\tau_N^1+\tau_N^2)}{\sqrt{N}} \right) = (0, H_{y,\overline{v}}, y, \overline{v}+\kappa_3 y H_{y,\overline{v}}), \tag{3.49}$$

holds, where

$$V_1 \stackrel{\text{def.}}{=} \overline{v}+\kappa_3 y H_{y,\overline{v}} \stackrel{\text{dist.}}{=} \sqrt{\overline{v}^2+2\frac{\kappa_3}{\kappa_4} y E_1}, \tag{3.50}$$

where E_1 is an exponentially distributed random variable with parameter 1 by Relation (3.44). Using again Skorohod's representation theorem, it can be assumed that, with a convenient probability space, the convergence (3.49) holds almost surely.

Let $v_1^N = X_4^N(t_1+\tau_N^1+\tau_N^2)$ and $(A_1^N(t))$ the solution of (3.42) associated to an independent Poisson process $\mathcal{P}_5$ and with initial point v_1^N. The convergence

$$\lim_{N\to+\infty} \left(\frac{A_1^N(t)}{\sqrt{N}}, 0\leq t<T \right) = \left(V_1 e^{-\kappa_5 t}, 0\leq t<T \right)$$

holds and, up to a change of probability space, it holds almost surely for the uniform topology.

Our time change $(\ell_0^N(t))$ "removes" the time interval $[t_1+\tau_N^1, t_1+\tau_N^1+\tau_N^2)$. We define

$$\begin{aligned}\left(\widetilde{X}_2^N(t), \widetilde{X}_4^N(t)\right) &= \mathbb{1}_{\left\{t<t_1+\tau_N^1\right\}} \left(y_N, A_0^N(t)\right) \\ &\qquad + \mathbb{1}_{\left\{t_1+\tau_N^1\leq t\right\}} \left(X_2^N(t_1+\tau_N^1+\tau_N^2), A_1^N(t-t_1-\tau_N^1)\right) \\ &= \left(y_N, A_0^N(t)\right) + \mathbb{1}_{\left\{t_1+\tau_N^1\leq t\right\}} \left(X_2^N(t_1+\tau_N^1+\tau_N^2)-y_N, A_1^N(t-t_1-\tau_N^1)-A_0^N(t)\right).\end{aligned}$$

Relation (3.47) gives that, for the J_1-topology,

$$\lim_{N\to+\infty} \left(\frac{\widetilde{X}_2^N(t)}{N}, \frac{\widetilde{X}_4^N(t)}{\sqrt{N}} \right) = (y, V_y(t)) \stackrel{\text{def.}}{=} \left(y, \mathbb{1}_{\{t<t_1\}} ve^{-\kappa_5 t} + \mathbb{1}_{\{t_1\leq t\}} V_1 e^{-\kappa_5(t-t_1)} \right).$$

Note that on the event $t_1 \leq T < t_2$, we have

$$\left(\frac{\widetilde{X}_2^N(t)}{N}, \frac{\widetilde{X}_4^N(t)}{\sqrt{N}} \right) \stackrel{\text{dist.}}{=} \left(\frac{X_2^N(\ell_0^N(t))}{N}, \frac{X_4^N(\ell_0^N(t))}{\sqrt{N}} \right)$$

if the second component of $(X_N(t))$ does not increase on $[t_1, t_1+\tau_N^1]$, i.e. on the event $\mathcal{E}_N = \left\{\mathcal{P}_1((0,\kappa_1)\times[t_1, t_1+\tau_N^1]) = 0\right\}$. The probability of $\mathcal{E}_N$ is arbitrarily close to 1 as N gets large since $(\sqrt{N}\tau_N^1)$ converges to 0. Hence if Φ is a bounded continuous functional on $\mathcal{D}([0,T])$ endowed with the J_1-topology, we obtain the relation

$$\lim_{N\to+\infty} \mathbb{E}\left(\Phi\left(\left(\frac{X_2^N(\ell_0^N(t))}{N}, \frac{X_4^N(\ell_0^N(t))}{\sqrt{N}} \right) \right) \mathbb{1}_{\{t_1\leq T<t_2\}} \right) = \mathbb{E}\left(\Phi\left((y, V_y(t)) \right) \mathbb{1}_{\{t_1\leq T<t_2\}} \right).$$

note that $V_y(t_1-)=\overline{v}$ and

$$V_y(t_1)-V_y(t_1-) = \sqrt{V_y(t_1-)^2+2\frac{\kappa_3}{\kappa_4} y E_1} - V_y(t_1-),$$

by Relation (3.50). We conclude that on the event $\{t_1 \leq T < t_2\}$, the processes

$$\left(\frac{X_2^N(\ell_0^N(t))}{N}, \frac{X_4^N(\ell_0^N(t))}{\sqrt{N}}, t \in [0,T] \right)$$

converge in distribution for the J_1-topology to $(V_y(t))$, which can be expressed as the solution of the SDE (3.33), with $\alpha=\kappa_0$, $\beta=\kappa_4/(2\kappa_3 y)$ and $\gamma=1/\kappa_5$. Recall that the points (t_n) do not depend on N.

It is straightforward, by induction, to extend this result. For any n, the relation

$$\lim_{N\to+\infty} \mathbb{E}\left(\Phi\left(\left(\frac{X_2^N(\ell_0^N(t))}{N}, \frac{X_4^N(\ell_0^N(t))}{\sqrt{N}} \right) \right) \mathbb{1}_{\{t_n \leq T < t_{n+1}\}} \right) = \mathbb{E}\left(\Phi\left((y, V_y(t)) \right) \mathbb{1}_{\{t_n \leq T < t_{n+1}\}} \right),$$

holds.

With the martingale problem formulation associated to the SDE (3.33), it is not difficult to see that $(V_y(t))$ is a Markov process with infinitesimal generator $\mathcal{A}_y$. See Section 4.4 of Ethier and Kurtz [24]. The proposition is proved. □

6.3. Convergence on the Normal Timescale. We can now state convergence results for the processes $(X_2^N(t)/N, X_4^N(t)/\sqrt{N})$ without a time-change. With the same notations as in the proof of the last proposition, one has to consider, for example, time intervals of the type $[t_1+\tau_N^1, t_1+\tau_N^1+\tau_N^2)$. By Proposition 3.13, its width τ_N^2 is converging in distribution to 0, whereas the number of jumps on it is of the order of $\sqrt{N}$. Because of that, the J_1-topology is not a convenient topology for a convergence result.

We will use another Skorohod topology, the M_1-topology, which allows the accumulation of jumps in a small neighborhood. We will also consider a non-Skorohod topology, the S-topology. Both topologies have their pros and cons, see [44]. For the S-topology, tightness criteria are somewhat simpler and the sum (of processes) is a continuous mapping for this topology, which is not the case for the M_1 and J_1-topologies. Unfortunately this is not a metrisable topology and the S convergence does not imply the convergence in distribution of the finite marginals. For both topologies, the integration of processes is a continuous functional, which is a key property to investigate the asymptotic integral equations verified by the possible limiting points.

Chapter 12 of Whitt [77] contains an in-depth presentation of the M_1-Skorohod topology with many details (and other Skorohod topologies) and Jakubowski [44] for the S-topology. See also Kern [49] for a quick and nice introduction to the intricacies of Skorohod topologies.

PROPOSITION 3.16. *If $X_N(0)=(0, y_N, 1, v_N)$ satisfy Relation (3.41), then for the convergence in distribution for the M_1-Skorohod topology and the S-topology,*

$$\lim_{N\to+\infty} \left(\frac{X_2^N(t)}{N}, \frac{X_4^N(t)}{\sqrt{N}} \right) = (y, V_y(t)),$$

where $(V_y(t))$ is the Markov process on $\mathbb{R}_+$ defined in Proposition 3.15.

PROOF. Recall that (t_n) is the non-decreasing sequence of points of $\mathcal{P}_0((0,\kappa_0], \mathrm{d}t)$, as in the proof of Proposition 3.15 they can be considered as "fixed" since they do not depend on N. We proceed as in the proof of Proposition 3.15 by working on the events $\{t_n \leq T < t_{n+1}\}$. As before, it is sufficient to consider $n=1$. As before, the Skorohod representation theorem is used, implicitly this time, with a convenient probability space, ...

We begin with the M_1-topology. Compared to the proof of the last proposition, the same type of arguments are used. We only have to "insert" the time interval

$[t_1+\tau_N^1, t_1+\tau_N^1+\tau_N^2)$ which will give the jump at time t_1. Recall that the M_1-topology is weaker than the J_1-topology.

We introduce the processes $(\widetilde{X}_2^N(t), \widetilde{X}_4^N(t))$ defined by

$$(\widetilde{X}_2^N(t), \widetilde{X}_4^N(t)) = (X_2^N(t), X_4^N(t)), \quad t\notin(t_1+\tau_N^1, t_1+\tau_N^1+\tau_N^2]$$

and, for $t\leq T$,

$$\begin{cases} \widetilde{X}_2^N(t)=X_2^N(t_1+\tau_N^1) \text{ for } t\in(t_1+\tau_N^1, t_1+\tau_N^1+\tau_N^2) \\ \widetilde{X}_2^N(t_1+\tau_N^1+\tau_N^2)=X_2^N(t_1+\tau_N^1+\tau_N^2) \end{cases}$$

and

$$\widetilde{X}_4^N(t) = X_4^N(t_1+\tau_N^1)+\kappa_3 y N(t-(t_1+\tau_N^1)) \text{ for } t\in(t_1+\tau_N^1, t_1+\tau_N^1+\tau_N^2).$$

By using the notations of the last proof, on the time interval $[t_1+\tau_N^1, t_1+\tau_N^1+\tau_N^2)$, $(\widetilde{X}_2^N(t))$ is kept constant equal to $X_2^N(t_1+\tau_N^1)$ and $(\widetilde{X}_4^N(t)/\sqrt{N})$ is a linear interpolation between the points

$$\frac{X_4^N(t_1+\tau_N^1)}{\sqrt{N}} \overset{\text{dist}}{\sim} \overline{v} = v\exp(-\kappa_5 t_1)$$

$$\text{and } \frac{X_4^N(t_1+\tau_N^1)}{\sqrt{N}}+\kappa_3 y\sqrt{N}\tau_2^N \overset{\text{dist}}{\sim} V_1=\overline{v}+\kappa_3 y H_{y,\overline{v}},$$

which correspond to $V_y(t_1-)$ and $V_y(t_1)$.

By using Corollary 3.14, where the underlying topology is the uniform norm, and Definition (3.4) of Whitt [77] for the distance for M_1 on $\mathcal{D}([0,T))$, on the event $\{t_1\leq T<t_2\}$, the two processes $(\widetilde{X}_2^N(t)/N, \widetilde{X}_4^N(t)/\sqrt{N})$ and $(X_2^N(t)/N, X_4^N(t)/\sqrt{N})$ are arbitrarily close as N goes to infinity.

The classical example (3.1) p. 80 of Whitt [77] and its parametrization (3.4) show that on the event $\{t_1\leq T<t_2\}$, the sequence of processes

$$\left(\frac{\widetilde{X}_2^N(t)}{N}, \frac{\widetilde{X}_4^N(t)}{\sqrt{N}}\right)$$

converges in distribution for the M_1-topology to the process $(y, V_y(t))$.

For the S-topology, for $x\in\mathcal{D}((0,T))$ and $\eta>0$, $N_\eta(x)$ is the number of oscillations of order η for $(x(t))$ on $[0,T]$, i.e. for $k\geq 1$, $N_\eta(x)\geq k$ holds if there exists $t_1<t_2<\cdots<t_{2k}$, such that $|x(t_{2i})-x(t_{2i-1})|>\eta$ for $i\in\{1,\ldots,k\}$.

Proposition (3.1) (iii) of Jakubowski [44] shows that if, for any $\eta>0$, the sequence of random variables

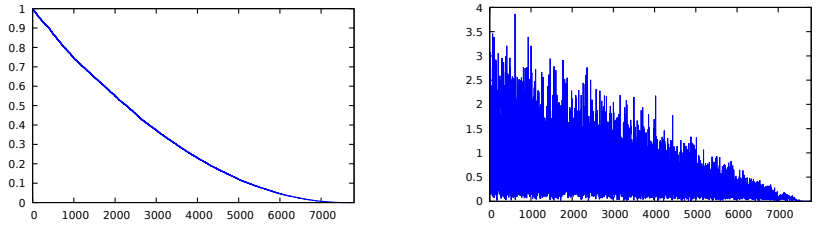
$$\left(\left\|\frac{X_2^N}{N}\right\|_T, \left\|\frac{X_4^N}{\sqrt{N}}\right\|_T, \frac{X_2^N(t)}{N}, N_\eta\left(\frac{X_2^N(t)}{N}\right), N_\eta\left(\frac{X_4^N(t)}{\sqrt{N}}\right)\right)$$

is tight, then the sequence of processes $(X_2^N(t)/N, X_4^N(t)/\sqrt{N})$ is tight for the S-topology. This is seen by the same arguments as in the proof of Proposition 3.15 for the time intervals $[0,t_1)$, $[t_1, t_1+\tau_1^N)$ and, with Corollary 3.14 for the time interval $[t_1+\tau_1^N, t_1+\tau_1^N+\tau_2^N)$.

Since, by the M_1-convergence, the finite marginals of these processes converge in distribution to the corresponding finite marginals of $(y, V_y(t))$, Theorem 3.11 of Jakubowski [44] gives the convergence in distribution for the S-topology to $(y, V_y(t))$

The proposition is proved.

□



(A) $(X_2^N(t)/N)$ (B) $(X_4^N(t)/\sqrt{N})$

FIGURE 3. CRN with Four Nodes. Full simulation.
Initial state $(0, N, 0, 0)$, κ_i=1, i=0,...,5 and N=$9{\cdot}10^6$.

7. A Stochastic Averaging Principle

Starting from the state $(0, y_N, 1, v_N)$ with $(y_N, v_N)\sim(yN, v\sqrt{N})$, Proposition 3.13 shows that, when the value of $(X_3^N(t))$ switches back to 0, the process $(X_4^N(t))$ has increased by an amount of the order of $\sqrt{N}$. This implies that the process $(X_2^N(t))$ has decreased by an amount also of the order of $\sqrt{N}$ at this instant. Recall that starting from a state for which the third coordinate is 0, it becomes 1 again after a duration of time stochastically lower bounded by an exponential distribution with parameter κ_0. Hence, to have a decay of the order of N for $(X_2^N(t))$, one needs a number of such cycles of the order of $\sqrt{N}$ and, therefore, this suggests that the "correct" timescale to observe a decay of $(X_2^N(t))$ is $(\sqrt{N}t)$. This is the main result of this section.

We introduce $(U_N(t))$=$(U_1^N(t), U_2^N(t), U_3^N(t), U_4^N(t))$, with $(U_1^N(t))$=(0), and the other coordinates are the solution of the SDE

$$\begin{cases} \mathrm{d}U_2^N(t) &= -\mathcal{P}_3\left(\left(0, \kappa_3 U_2^N U_3^N(t-)\right), \mathrm{d}t\right), \\ \mathrm{d}U_3^N(t) &= \mathcal{P}_0\left((0,\kappa_0), \mathrm{d}t\right) - \mathcal{P}_4\left(\left(0, \kappa_4 U_3^N U_4^N(t-)\right), \mathrm{d}t\right), \\ \mathrm{d}U_4^N(t) &= \mathcal{P}_3((0, \kappa_3 U_2^N U_3^N(t-)), \mathrm{d}t) - \mathcal{P}_5((0, \kappa_5 U_4^N(t-)), \mathrm{d}t), \end{cases} \tag{3.51}$$

with $(U_N(0))$=$(0, y_N, 1, v_N)$ and the sequence (y_N, v_N) satisfies Relation (3.41).

We first give a heuristic motivation for the introduction of $(U_N(t))$. The difference between the processes $(X_N(t))$ and $(U_N(t))$ lies in the fact that when there is a jump of $\mathcal{P}_0\left((0,\kappa_0), \mathrm{d}t\right)$, for U_N it is transferred right away to the coordinate U_3^N and for X_N it goes to X_1^N and then to X_3^N at a rate $\kappa_2 X_2^N(t)$ at time t. If $X_2^N(t)$ is of the order of N, the difference between $(X_N(t))$ and $(U_N(t))$ holds on a time interval whose duration is of the order of $1/N$ and the probability that there is another event that will change the coordinates during that time is of the order of $1/N$. Hence on a time interval $[0, \sqrt{N}T]$, given that the number of jumps of $\mathcal{P}_0$ is of the order of $\sqrt{N}$, the two processes should be "close". Proposition 3.27 at the end of this section gives a formal assessment of this non-rigorous picture. The limit results for $(X_N(t))$ will be a consequence of the limit results obtained for $(U_N(t))$.

From now on, we investigate the asymptotic behavior of $(U_N(t))$. The general strategy is of first considering, via a time change, the time evolution of the process when the third coordinate is above 1

7.1. Time Evolution when U_3^N is non-zero.

We introduce the local time of the excursions of the process $(U_3^N(t))$ above 1.

DEFINITION 3.17 (Time Change). *For $t\geq 0$,*

$$L_1^N(t) \stackrel{\text{def.}}{=} \int_0^t \mathbb{1}_{\{U_3^N(s)\geq 1\}}\,\mathrm{d}s \quad \textit{and} \quad \ell_1^N(t) \stackrel{\text{def.}}{=} \inf\{s{>}0 : L_1^N(s){>}t\}.$$

Before introducing the formal time change arguments of this section, to motivate the description of the process $(U_N(t))$ on the timescale $(\ell_1^N(t))$, we give a quick presentation of the dynamics involved.

In view of the SDEs (3.51), if U_3^N is 1 at time t_0, it becomes 0 at rate $\kappa_4 U_4^N(t_0-)$ and stays at 0 for an exponentially distributed amount of time with parameter κ_0.

When $(U_3^N(t))$ is 0 on a time interval $[t_0, t_0+a)$, for some $a>0$, the process $(U_2^N(t))$ does not change and the process $(U_4^N(t))$ can only decrease. At time t_0, the $p=U_4^N(t_0)$ elements of $(U_4^N(t))$ can be seen as having lifetimes $E_{\kappa_5,1}$, …, $E_{\kappa_5,p}$. These variables are exponentially distributed with parameter κ_5. At time t_0+a there remain only those whose lifetime is greater than a.

This suggests quite naturally the following description of our system. Let $\mathcal{N}$ be a Poisson marked point process on $\mathbb{R}_+^2\times\mathbb{R}_+\times\mathbb{R}_+^{\mathbb{N}}$ with intensity measure

$$\mathrm{d}s\otimes\mathrm{d}t\otimes\kappa_0\exp(-\kappa_0 a)\,\mathrm{d}a\otimes Q(\mathrm{d}b),$$

where Q is the distribution of $(E_{\kappa_5,i})$ on $\mathbb{R}_+^{\mathbb{N}}$.

The process $(Z_N(t))=(Z_2^N(t), Z_3^N(t), Z_4^N(t))$ is the solution of the SDE

$$\begin{cases} \mathrm{d}Z_2^N(t) = -\mathcal{P}_3\left(\left(0,\kappa_3 Z_2^N Z_3^N(t-)\right),\mathrm{d}t\right), \\ \mathrm{d}Z_3^N(t) = \mathcal{P}_0\left((0,\kappa_0),\mathrm{d}t\right) - \mathbb{1}_{\left\{Z_3^N(t-)\geq 2\right\}}\mathcal{P}_4\left(\left(0,\kappa_4 Z_3^N Z_4^N(t-)\right),\mathrm{d}t\right), \\ \mathrm{d}Z_4^N(t) = \mathcal{P}_3((0,\kappa_3 Z_2^N Z_3^N(t-)),\mathrm{d}t) - \mathcal{P}_5((0,\kappa_5 Z_4^N(t-)),\mathrm{d}t), \\ \qquad -\mathbb{1}_{\left\{Z_3^N(t-)=1\right\}}\displaystyle\int_{a,b} S(Z_4^N(t-),a,b)\mathcal{N}\left(\left(0,\kappa_4 Z_4^N(t-)\right],\mathrm{d}t,\mathrm{d}a,\mathrm{d}b\right), \end{cases} \tag{3.52}$$

where $S(\cdot)$ is defined by Relation (3.35). It turns out that $(Z_N(t))$ has the same distribution as the time changed process $(U_N(t))$.

Proposition 3.18. *For $N\geq 1$, if $(U_N(0))=(Z_N(0))=(y_N,1,v_N)$ then*

$$(Z_N(t)) \overset{\text{dist.}}{=} \left(U_N(\ell_1^N(t))\right),$$

where $(U_N(t))$, $(Z_N(t))$, and $(\ell_1(t))$ are respectively defined by Relations (3.51), (3.52) *and Definition* (3.17)

Proof. We give a sketch of the proof. It is essentially a consequence of a repeated use of the strong Markov property of Poisson processes. If $U_N(0)=(y_N,1,v_N)$, let τ_U^N be the first instant when $(U_3^N(t))$ hits 0, i.e. the first jump of the counting process

$$\left(\int_0^t \mathbb{1}_{\left\{U_3^N(s)=1\right\}}\mathcal{P}_4\left(\left(0,\kappa_4 U_4^N(s-)\right),\mathrm{d}s\right)\right).$$

On the time interval $[0,\tau_U^N)$ $(U_N(t))$ satisfies the SDEs (3.52). If E_{κ_0} is the first point of $(\mathcal{P}_0(\left(0,\kappa_0\right),(\tau_U^N,\tau_U^N+t])$, E_{κ_0} is an exponential random variable with parameter κ_0 independent of $\mathcal{F}_{\tau_U^N}$, and on the time interval $[\tau_U^N,\tau_U^N+E_{\kappa_0})$, the coordinate $(U_4^N(\tau_U^N+t))$ satisfies the SDE,

$$\mathrm{d}A_N(t) = -\mathcal{P}_5((0,\kappa_5 A_N(t-)),\tau_U^N+\mathrm{d}t),$$

with $A_N(0)=U_4^N(\tau_U^N)$. It is easily seen that the process $(A_N(t))$ has the same distribution as

$$\left(\sum_{i=1}^{U_4^N(\tau_U^N)} \mathbb{1}_{\left\{E_{\kappa_5,i}>t\right\}}\right),$$

where $(E_{\kappa_5,i})$ are i.i.d. exponential random variables with parameter κ_5 independent of $\mathcal{F}_{\tau_U^N}$. We have therefore that

$$U_4(\ell_1^N(\tau_U^N))-U_4(\ell_1^N(\tau_U^N-)) = U_4(\tau_U^N+E_{\kappa_0})-U_4(\tau_U^N)$$
$$= \sum_{i=1}^{U_4^N(\tau_U^N)} \mathbb{1}_{\{E_{\kappa_5,i}>E_{\kappa_0}\}} - U_4^N(\tau_U^N) = -S(U_4^N(\tau_U^N), E_{\kappa_0}, (E_{\kappa_5,i})),$$

where S is the function defined by Relation (3.35). One can proceed by induction on the successive instants of return to 1 from 0 of $(U_3^N(t))$. The proposition is proved. □

Note that by considering the process $(Z_N(t))$, we remove the time intervals where $(U_2^N(t)/N)$ is constant and therefor e no effect on its decay. The relevant timescale to see the decrease of $(Z_2^N(t)/N)$ is now the normal timescale (t), and we will see later that $\ell_1^N(t)=O(\sqrt{N})$.

Definition 3.19. *The occupation measure m_N of $(Z_4^N(t)/\sqrt{N})$ is*

$$\langle m_N, F\rangle = \int_0^{+\infty} F\left(s, \frac{Z_4^N(s)}{\sqrt{N}}\right) \mathrm{d}s,$$

for $F\in\mathcal{C}_c(\mathbb{R}_+^2)$.

We can now state a key result of this section.

Theorem 3.20. *If $(Z_N(t))$, the solution of the SDEs (3.52) and μ_Z^N the occupation measure of Definition 3.19 are such that $(Z_N(0))=(y_N,1,v_N)$ and (y_N,v_N) satisfies Relation (3.41), then, for the convergence in distribution,*

$$\lim_{N\to+\infty}((Z_2^N(t)),\mu_Z^N) = ((z_2(t)),\mu_Z^\infty)\,,$$

with $(z_2(t))=(ye^{-\kappa_3 t})$, and, for any $F\in\mathcal{C}_c(\mathbb{R}_+^2)$,

$$\int_{\mathbb{R}_+^2} F\,(s,x)\,\mu_Z^\infty(\mathrm{d}s,\mathrm{d}x) = \int_{\mathbb{R}_+^2} F(s,\sqrt{x})\Gamma_0\left(\frac{\kappa_0}{2\kappa_5}+\frac{1}{2},\frac{\kappa_4}{2\kappa_3 z_2(s)}\right)(\mathrm{d}x)\,\mathrm{d}s,$$

where $\Gamma_0(\cdot,\cdot)$ is the distribution of Definition 3.9.

The proof of this theorem is carried out dividing the time integral according to the value of $Z_3^N(t)$, recall that it is always above 1.

(a) In the time intervals where $Z_3^N(\cdot)=(1)$, we will show that the limit result can be deduced from Theorem 3.12 of Section 5.3.

(b) For the time intervals during which $Z_3^N(t)\geq 2$, the goal will be of showing that these intervals do not contribute to the final limit, more precisely that, for the convergence in distribution,

$$\lim_{N\to+\infty}\left(\int_0^{+\infty}\mathbb{1}_{\{Z_3^N(s)\geq 2\}}F\left(s,\frac{Z_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s\right)=0,$$

for some convenient class of functions on $\mathbb{R}_+^2$. This is done in the following way.

If t_1^3 is such that $Z_3^N(t_1^3)=2$, we define $\tau_Z^N=\inf\{t\geq 0 : Z_3^N(t_3^N+t)=1\}$. We will show that $\tau_Z^N=O(1/\sqrt{N})$,

— first, we show after a time of the order of $O(1/\sqrt{N})$, the process $(Z_4^N(t_1^3+t))$ reaches a value the of order $\sqrt{N}$, and stays at this order of magnitude.

— secondly, since $Z_3^N(t)$ decreases at a rate $\kappa_4 Z_4^N(t)$, we show that $Z_3^N(t)$ goes back to 1 after a time of the order of $O(1/\sqrt{N})$.

And to conclude, we only have to notice that these time intervals $(t_1^3, t_1^3 + \tau_Z^N)$ are in a finite number in $[0,T]$ for any $T \geq 0$, since every time interval is separated by an exponentially distributed variable with parameter κ_0.

The proof will be done by induction, on an "excursion" of $Z_3^N(t)$ over 2. On a first time interval $[t_1^3, t_1^3+\tau_Z^N)$, and then iterating.

PROOF. If $(Z_N(0))=(y_N, 1, v_N)$, t_1^3, the first instant of jump $+1$ of $(Z_3^N(t))$ has an exponential distribution with parameter κ_0. Up to time t_1^3 the process $(Z_2^N(t), Z_4^N(t))$ has the same distribution as the process $(A_N(t), B_N(t))$, the solution of the SDE,

$$\begin{cases} \mathrm{d}A_N(t) = -\mathcal{P}_3\left((0, \kappa_3 A_N(t-)\right), \mathrm{d}t\right), \\ \mathrm{d}B_N(t) = \mathcal{P}_3((0, \kappa_3 A_N(t-)), \mathrm{d}t) - \mathcal{P}_5((0, \kappa_5 B_N(t-)), \mathrm{d}t), \\ \qquad - \displaystyle\int_{a,b} S(B_N(t-), a, b)\mathcal{N}\left((0, \kappa_4 B_N(t-)\right], \mathrm{d}t, \mathrm{d}a, \mathrm{d}b). \end{cases} \tag{3.53}$$

with $(A_N(0), B_N(0))=(y_N, v_N)$.

It is straightforward to show that $(A_N(t)/N)$ is converging in distribution to $(y\exp(-\kappa_3 t))$ and that the asymptotic behavior of the occupation measure associated to $(B_N^0(t)/\sqrt{N})$ can be obtained from Theorem 3.12 of Section 5.3. It is not difficult to see that the additional term due to $\mathcal{P}_5$ in the SDE defining $(Z_4^N(t))$ does not play a role for this limit result.

If $Z_3^N(t_1^3)=2$, let τ_Z^N , be the hitting time of 1 of the process $(Z_3^N(t+t_1^3))$. In view of the first equation of SDE (3.52), with high probability, we have that, for any $t\in[0,T]$, $Z_2^N(t)\geq\delta N$. To simplify our arguments, since we are dealing with convergence in distribution, we assume that this relation holds almost surely, it is not difficult to modify our proof accordingly, at the expense of additional terms. If $(C_N(t), D_N(t))$ is the solution of the SDE,

$$\begin{aligned} \mathrm{d}C_N(t) &= \widetilde{\mathcal{P}}_0\left((0, \kappa_0), \mathrm{d}t\right) - \mathbb{1}_{\{C_N(t-)>0\}}\widetilde{\mathcal{P}}_4\left((0, \kappa_4(1+C_N(t-))D_N(t-)\right), \mathrm{d}t\right), \\ \mathrm{d}D_N(t) &= \widetilde{\mathcal{P}}_3((0, 2\kappa_3\delta N), \mathrm{d}t) - \widetilde{\mathcal{P}}_5((0, \kappa_5 D_N(t-)), \mathrm{d}t), \end{aligned}$$

with $C_N(0)=1$ and $D_N(0)=0$, and where, for $i\in\{0,3,4,5\}$, $\widetilde{\mathcal{P}}_i(\mathrm{d}s, \mathrm{d}t)$ is the Poisson process $\mathcal{P}_i$ shifted at t_1^3, i.e. $\mathcal{P}_i(\mathrm{d}s, t_1^3+\mathrm{d}t)$. A simple coupling can be constructed so that the relations $D_N(t)\leq Z_4^N(t_1^3+t)$ and $Z_3^N(t)\leq C_N(t_1^3+t)+1$ hold for all $0\leq t\leq\tau_Z^N$.

Standard arguments of stochastic calculus, give that, for the convergence in distribution, the relation

$$\lim_{N\to+\infty}\left(\frac{D_N(t/\sqrt{N})}{\sqrt{N}}\right) = (2\kappa_3\delta t)$$

holds and if τ_D^N is the hitting time of $\lceil\kappa_3\delta\sqrt{N}\rceil$, then

$$\limsup_{N\to+\infty}\sqrt{N}\mathbb{E}(\tau_D^N) < +\infty,$$

and, since $\left(C_N(\tau_D^N)-1\right)^+$ is bounded by the number of new arrivals for $(C_N(t))$ on the time interval $[0, \tau_D^N)$

$$\limsup_{N\to+\infty}\sqrt{N}\mathbb{E}\left(\left(C_N(\tau_D^N)-1\right)^+\right) < +\infty.$$

For $T>0$, starting from τ_D^N, the process $(D_N(t))$ stays above $\kappa_3\delta\sqrt{N}$ with high probability on a time interval $[\tau_D^N, \tau_D^N+T/\sqrt{N}]$. With the same argument as before, we assume that this relation holds almost surely.

If τ_C^N is the hitting time of 0 by $(C_N(\tau_D^N+t))$, then $\tau_Z^N \le \tau_D^N+\tau_C^N$. The integration of the SDE for $(C_N(t))$ gives the relation

$$\mathbb{E}\left(C_N(\tau_D^N)\right)+\kappa_0\mathbb{E}(\tau_C^N\wedge t)-\kappa_4\kappa_3\delta\sqrt{N}\mathbb{E}\left(\int_0^{\tau_C^N\wedge t}(1+C_N(\tau_D^N+s))\,\mathrm{d}s\right) = \mathbb{E}\left(C_N(\tau_D^N+\tau_C^N\wedge t)\right),$$

hence, since $C_N(t)\ge 1$ on the time interval $[\tau_D^N, \tau_D^N+\tau_C^N)$, we obtain

$$\left(\kappa_4\kappa_3\delta\sqrt{N}-\kappa_0\right)\mathbb{E}\left(\tau_C^N\wedge t\right) \le \mathbb{E}\left(C_N(\tau_D^N)\right) \le 1+\mathbb{E}\left(\left(C_N(\tau_D^N)-1\right)^+\right),$$

hence

$$\limsup_{N\to+\infty}\sqrt{N}\mathbb{E}(\tau_C^N) < +\infty.$$

By gathering these results, and since, for $t\ge 0$,

$$Z_4^N(t_1^3+t/\sqrt{N}) - Z_4^N(t_1^3) \le \widetilde{\mathcal{P}}_3((0,2\kappa_3\delta N),[0,t/\sqrt{N}]),$$

we have finally obtained that

$$\limsup_{N\to+\infty}\sqrt{N}\mathbb{E}(\tau_Z^N) < +\infty, \quad \text{and} \quad \limsup_{N\to+\infty}\frac{\mathbb{E}(Z_4^N(t_1^3+\tau_Z^N)-Z_4^N(t_1^3))}{\sqrt{N}} < +\infty.$$

This shows that, if f is a bounded Borelian function on $\mathbb{R}_+^2$, the sequences of random variables

$$\left(\int_{t_1^3}^{t_1^3+\tau_Z^N}\mathbb{1}_{\{Z_3^N(s)\ge 2\}}f\left(\frac{Z_2^N(s)}{N},\frac{Z_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s\right) \text{ and } \left(\left|\frac{Z_2^N(t_1^3+\tau_Z^N)-Z_2^N(t_1^3)}{N}\right|\right)$$

converge in distribution to 0. We can now apply the convergence result, Theorem 3.12, for the processes $(Z_2^N(t), Z_4^N(t))$ starting from time $t_1^3+\tau_Z^N$, as if it was starting from time t_1^3 with $Z_3^N(t_1^3)=1$, since their first coordinate is of the same order of magnitude in N, and that the limit result for the occupation measure does not depend on the initial value of $Z_4^N(t_1^3+\tau_Z^N)$ as long as it is of the order of N at most.

We conclude the proof of the proposition by induction on the successive jumps (t_i^3) of $\mathcal{P}_0((0,\kappa_0),\mathrm{d}t)$ on the time interval $[0,T]$. □

Note that the induction holds because the averaging result of Theorem 3.12 do not need for $Z_4^N(t)$ to start from a state of the order of $\sqrt{N}$, but only $O(N)$. This particular property is vital here, since we do not have the control of the process $(Z_4^N(t)/\sqrt{N})$ over some time interval, but only of its time integrals.

We have seen in the proof that the process $(Z_3^N(t))$ is identically 1, *as long as we consider time integrals*. The following Corollary is easily deduced from Relations (3.37) and (3.53).

COROLLARY 3.21. *Let $F\in\mathcal{C}(\mathbb{R}_+^2)$ with support on $[0,T]\times\mathbb{R}_+$ such that there exist c_1, $c_2>0$ such that for any $s,x\in\mathbb{R}_+$, $F(s,x)\le c_1+c_2x$ holds for $(s,x)\in[0,T]\times\mathbb{R}_+$, then under the assumptions of Theorem 3.20, the sequence*

$$\left(\int_0^{+\infty}\mathbb{1}_{\{Z_3^N(s)=1\}}F\left(s,\frac{Z_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s\right)$$

converges in distribution to $\langle\mu_Z^\infty, F\rangle$.

7.2. Time Evolution on the Timescale $(\sqrt{N}t)$. We are now going to express $(U_N(t))$ in terms of $(Z_N(t))$. Recall that the process $(Z_2^N(t), Z_4^N(t))$ analyzed in the last section is just the process $(U_2^N(t), U_4^N(t))$ with the time intervals during which $(U_3^N(t))$ is 0 removed. See Proposition 3.18.

We denote by $(Z_N(t))$ the solution of the SDE (3.52) starting from $(y_N, 1, v_N)$. We define

(3.54)
$$(H_N(t)) = \left(t + \int_{(0,t]\times\mathbb{R}_+\times\mathbb{R}_+^{\mathbb{N}}} a \mathbb{1}_{\left\{Z_3^N(s-)=1\right\}} \mathcal{N}\left(\left(0, \kappa_4 Z_4^N(s-)\right], \mathrm{d}s, \mathrm{d}a, \mathrm{d}b\right)\right),$$

and the (potential) hitting time of 0 for $(Z_3^N(t))$ is defined as τ_N, i.e.

$$\tau_N = \inf\left\{ t : \int_{(0,t]\times\mathbb{R}_+\times\mathbb{R}_+^{\mathbb{N}}} \mathbb{1}_{\left\{Z_3^N(s-)=1\right\}} \mathcal{N}((0, \kappa_4 Z_4^N(s-)], \mathrm{d}s, \mathrm{d}a, \mathrm{d}b) \neq 0 \right\}.$$

Strictly speaking, this is an incorrect presentation for τ_N since $(Z_3^N(t))$ never visits 0. This is in fact meant for $(U_3^N(t))$, as long as the two processes coincide. Additionally, $(a_N, b_N) \in \mathbb{R}_+ \times \mathbb{R}_+^{\mathbb{N}}$ is the mark associated to τ_N, i.e.

$$\mathcal{N}((0, \kappa_4 Z_4^N(\tau_N-)], \{\tau_N\}, \mathrm{d}a, \mathrm{d}b) = \delta_{(a_N, b_N)},$$

a_N and $b_N = (b_{N,i})$ are independent and independent of $\mathcal{F}_{\tau_n}$, with respective distributions, an exponential law with parameter κ_0 and the law of an i.i.d. sequence of exponential random variables with parameter κ_5. We define $t_N = \tau_N + a_N$.

We now construct a process $(\widetilde{U}_N(t)) = (\widetilde{U}_2^N(t), \widetilde{U}_3^N(t), \widetilde{U}_4^N(t))$ with initial state $(y_N, 1, v_N)$ and

(3.55)
$$\begin{cases} t < \tau_N, & (\widetilde{U}_2^N, \widetilde{U}_3^N, \widetilde{U}_4^N)(t) = (Z_2^N, Z_3^N, Z_4^N)(t), \\ \tau_N \le t < t_N, & \begin{cases} (\widetilde{U}_2^N, \widetilde{U}_3^N)(t) = (Z_2^N(\tau_N-), 0), \\ \widetilde{U}_4^N(t) = \sum_{i=1}^{Z_4^N(\tau_N-)} \mathbb{1}_{\left\{b_{N,i} > t-\tau_N\right\}} = Z_4^N(\tau_N-) - S\left(Z_4^N(\tau_N-), t-\tau_N, b_N\right), \end{cases} \\ t = t_N, & (\widetilde{U}_2^N, \widetilde{U}_3^N, \widetilde{U}_4^N)(t_N) = \left(\widetilde{U}_2^N(t_N-), 1, Z_4^N(\tau_N-) - S\left(Z_4^N(\tau_N-), a_N, b_N\right)\right). \end{cases}$$

We have constructed the process $(\widetilde{U}_N(t))$ between two visits of the third coordinate to 1. We can construct by induction the process on the whole real half-line.

PROPOSITION 3.22. *For $N \ge 1$, the processes $(U_N(t))$ and $(\widetilde{U}_N(t))$ defined respectively by Relations* (3.51) *and* (3.55) *have the same distribution. Furthermore the relations*

$$\left(\int_0^{H_N(t)} \mathbb{1}_{\left\{\widetilde{U}_3^N(s) \ge 1\right\}} \mathrm{d}s \right) = (t) \text{ and } \left(\widetilde{U}_2^N\left(H_N(t)\right) \right) = \left(Z_2^N(t) \right)$$

hold.

We have in particular the identity $(H_N(t)) \overset{\text{dist.}}{=} (\ell_1^N(t))$. See Definition 3.17.

PROOF. The proof of the identity in distribution is analogous to the proof of Proposition 3.18. It relies again on strong Markov properties of Poisson processes and the representation of the process $(A_N(t))$ used in this proof.

The first relation comes directly from the construction of $(\widetilde{U}_N(t))$. The last relation is a consequence of the fact that $(\widetilde{U}_2^N(t))$ does not change on the times intervals where $(\widetilde{U}_3^N(t))$ is null. □

PROPOSITION 3.23. *If $(Z_N(t))$ is the solution of the SDEs* (3.52) *with the initial condition $(y_N, 1, v_N)$ and (y_N, v_N) satisfying Relation* (3.41) *then, for the convergence in distribution,*

$$\lim_{N\to+\infty}\left(\frac{H_N(t)}{\sqrt{N}}\right)=\left(t_\infty\left(1-e^{-\kappa_3 t/2}\right)\right),$$

with

$$t_\infty \stackrel{\text{def.}}{=} \sqrt{y}\frac{\sqrt{2\kappa_4}}{\kappa_5\sqrt{\kappa_3}}\frac{\Gamma(\kappa_0/(2\kappa_5))}{\Gamma(\kappa_0/(2\kappa_5)+1/2)}. \tag{3.56}$$

PROOF. For $t\geq 0$,

$$\frac{H_N(t)}{\sqrt{N}}=\frac{t}{\sqrt{N}}+M_N(t)+\frac{\kappa_4}{\kappa_0}\int_0^t \mathbb{1}_{\{Z_3^N(s-)=1\}}\frac{Z_4^N(s)}{\sqrt{N}}\,\mathrm{d}s,$$

where $(M_N(t))$ is a martingale whose previsible increasing process is

$$\left(\langle M_N\rangle(t)\right)=\left(2\frac{\kappa_4}{\kappa_0}\int_0^t \mathbb{1}_{\{Z_3^N(s-)=1\}}\frac{Z_4^N(s)}{N}\,\mathrm{d}s\right).$$

Theorem 3.20 and Doob's Inequality give that $(M_N(t))$ is converging in distribution to 0 and, with Corollary 3.21, that $(H_N(t)/\sqrt{N})$ converges to

$$\left(\frac{\kappa_4}{\kappa_0}\int_0^t\int_0^{+\infty}\sqrt{x}\Gamma\left(\frac{\kappa_0}{2\kappa_5}+\frac{1}{2},\frac{\kappa_4}{2\kappa_3 z_2(s)}\right)(\mathrm{d}x)\,\mathrm{d}s\right)$$
$$=\left(\frac{\kappa_4}{\kappa_0}\frac{\Gamma(\kappa_0/(2\kappa_5)+1)}{\Gamma((\kappa_0)/(2\kappa_5)+1/2)}\sqrt{2}\sqrt{\frac{\kappa_3}{\kappa_4}}\int_0^t\sqrt{z_2(s)}\,\mathrm{d}s\right),$$

by Relation (3.28), with $(z_2(t))=(y\exp(-\kappa_3 t))$. The proposition is proved. □

THEOREM 3.24. *If $(U_N(t))$ is the solution of the SDEs* (3.51) *with the initial condition $(y_N, 1, v_N)$ and (y_N, v_N) satisfying Relation* (3.41) *then, for the convergence in distribution,*

$$\lim_{N\to+\infty}\left(\frac{U_2^N\left(\sqrt{N}t\right)}{N},t<t_\infty\right)=(u_2(t))\stackrel{\text{def.}}{=}\left(y\left(1-\frac{t}{t_\infty}\right)^2,t<t_\infty\right)$$

with t_∞ is defined in Proposition 3.23.

The quadratic decay can be seen in the simulations of Figure 3 (A).

PROOF. For $0\leq t<t_\infty$, define

$$s_N(t)=\inf\{u>0 : H_N(u)>\sqrt{N}t\},$$

Proposition 3.23 gives the convergence in distribution

$$\lim_{N\to+\infty}(s_N(t))=\left(-\frac{2}{\kappa_3}\ln\left(1-\frac{t}{t_\infty}\right)\right). \tag{3.57}$$

We have the identities

$$\left(\frac{\widetilde{U}_2^N(\sqrt{N}t)}{N}\right)=\left(\frac{\widetilde{U}_2^N(H_N(s_N(t)))}{N}\right)=\left(\frac{Z_2^N((s_N(t)))}{N}\right),$$

the first one is due to the fact that $\widetilde{U}_2^N$ does not change just after a jump of $(H_N(t))$ and the second holds by Proposition 3.23.

We conclude the proof of the theorem with the convergence in distribution of Theorem 3.20 and Theorem 3.3 of Ethier and Kurtz [24]. □

DEFINITION 3.25. *The occupation measure* μ_U^N *of*

$$\left(\frac{U_4^N(\sqrt{N}t)}{\sqrt{N}}\right) \text{ is } \left\langle \mu_U^N, F\right\rangle \stackrel{\text{def.}}{=} \int_0^{+\infty} F\left(s, \frac{U_4^N(s\sqrt{N})}{\sqrt{N}}\right) \mathrm{d}s,$$

for $F \in \mathcal{C}_c(\mathbb{R}_+^2)$.

THEOREM 3.26. *If* $(U_N(t))$ *is the solution of the SDEs* (3.51) *with the initial condition* $(y_N, 1, v_N)$ *and* (y_N, v_N) *satisfying Relation* (3.41)*, then the sequence* (μ_U^N) *converges in distribution to* μ_U^∞ *defined by, for any* $F \in \mathcal{C}_c([0,t_\infty)\times\mathbb{R}_+)$,

$$\int F(s,x)\mu_U^\infty(\mathrm{d}s,\mathrm{d}x) = \int F(s,\sqrt{x})\Gamma_0\left(\frac{\kappa_0}{2\kappa_5}, \frac{\kappa_4}{2\kappa_3 y(1-t/t_\infty)^2}\right)(\mathrm{d}x)\,\mathrm{d}s,$$

where t_∞ *is defined by Relation* (3.56) *and* Γ_0 *is the distribution of Definition 3.9.*

By using Proposition 3.11 of Section 5.2, note that, for $0 \le t < t_\infty$,

$$\Gamma_0\left(\frac{\kappa_0}{2\kappa_5}, \frac{\kappa_4}{2\kappa_3 y(1-t/t_\infty)^2}\right)$$

is the invariant distribution of the infinitesimal generator $\mathcal{A}_{u_2(t)}$ of Proposition 3.15, where $(u_2(t))$ is defined by Theorem 3.24.

PROOF. Let $f \in \mathcal{C}_c([0,t_\infty)\times\mathbb{R}_+)$, we have

$$\left\langle \mu_U^N, F\right\rangle = \frac{1}{\sqrt{N}}\int_0^{+\infty} f\left(\frac{s}{\sqrt{N}}, \frac{U_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s.$$

For $F \in \mathcal{C}_c(\mathbb{R}_+)$ and $S \ge 0$, by using the definition of $(\widetilde{U}_4^N(s))$, the relation

$$\begin{aligned}(3.58)\quad \frac{1}{\sqrt{N}}\int_0^{H_N(S)} F\left(\frac{\widetilde{U}_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s &= \frac{1}{\sqrt{N}}\int_0^S F\left(\frac{Z_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s \\ + \frac{1}{\sqrt{N}}\int_0^S \mathbb{1}_{\{Z_3^N(s-)=1\}} \int_{(a,b)\in\mathbb{R}_+\times\mathbb{R}_+^{\mathbb{N}}} &\left(\int_0^a F\left(\frac{Z_4^N(s-)-S(Z_4^N(s-),u,b)}{\sqrt{N}}\right)\mathrm{d}u\right) \\ &\times \mathcal{N}((0,\kappa_4 Z_4^N(s-)],\mathrm{d}s,\mathrm{d}a,\mathrm{d}b),\end{aligned}$$

holds on the event $\{H_N(S) < \sqrt{N}t_\infty\}$ whose probability is converging to 1 as N gets large.

TIGHTNESS.
For $\varepsilon > 0$ and $0 \le t_0 < t_\infty$, with Proposition 3.23, we get that there exists $S > 0$ such that, for N sufficiently large

$$\mathbb{P}\left(\frac{H_N(S)}{\sqrt{N}} \notin [t_0, t_\infty)\right) \le \varepsilon,$$

hence

$$\begin{aligned}\mathbb{E}\left(\left\langle \mu_U^N, [0,t_0]\times[K,+\infty]\right\rangle\right) &= \frac{1}{\sqrt{N}}\int_0^{\sqrt{N}t_0} \mathbb{P}\left(U_4^N(s) \ge K\sqrt{N}\right)\mathrm{d}s \\ &\le \varepsilon t_0 + \frac{1}{\sqrt{N}}\mathbb{E}\left(\int_0^{H_N(S)} \mathbb{1}_{\left\{U_4^N(s)\ge K\sqrt{N}\right\}}\,\mathrm{d}s\right).\end{aligned}$$

With Relation (3.58), we obtain that

$$\begin{aligned}\frac{1}{\sqrt{N}}\mathbb{E}\left(\int_0^{H_N(S)} \mathbb{1}_{\left\{U_4^N(s)\geq K\sqrt{N}\right\}}\,\mathrm{d}s\right) &\leq \frac{1}{\sqrt{N}}\int_0^S \mathbb{P}\left(\frac{Z_4^N(s)}{\sqrt{N}}\geq K\right)\mathrm{d}s \\ &+\kappa_4\int_0^S \mathbb{E}\left(\mathbb{1}_{\left\{Z_3^N(s-)=1\right\}}\frac{Z_4^N(s)}{\sqrt{N}}\mathbb{1}_{\left\{Z_4^N(s)\geq K\sqrt{N}\right\}}\right)\mathrm{d}s \\ &\leq \frac{S}{\sqrt{N}}+\frac{1}{K}\int_0^S \mathbb{E}\left(\mathbb{1}_{\left\{Z_3^N(s-)=1\right\}}\left(\frac{Z_4^N(s)}{\sqrt{N}}\right)^2\right)\mathrm{d}s\end{aligned}$$

holds.

Relations (3.37) and (3.53) show that

$$\limsup_{N\to+\infty}\mathbb{E}\left(\int_0^S \mathbb{1}_{\left\{Z_3^N(s-)=1\right\}}\left(\frac{Z_4^N(s)}{\sqrt{N}}\right)^2\mathrm{d}s\right)<+\infty.$$

Therefore, one can choose K sufficiently large so that the quantity

$$\mathbb{E}\left(\left\langle\mu_U^N,[0,t_0]\times[K,+\infty]\right\rangle\right)$$

is arbitrarily small for N sufficiently large. Lemma 1.3 of Kurtz [55] gives the tightness of the sequence (μ_U^N) of random measures on $[0,t_\infty]\times\mathbb{R}_+$.

Identification of the Limit.

With the tightness property and since any limiting point can be represented as in Relation (3.14), it is enough to identify the limit of

$$\frac{1}{\sqrt{N}}\int_0^{\sqrt{N}t} F\left(\frac{\widetilde{U}_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s,$$

for any $t\in[0,t_\infty)$. Since, by Proposition 3.23, the process $(H_N(t)/\sqrt{N})$ converges in distribution to a deterministic function, one has to obtain the limit of the sequence

$$\left(\frac{1}{\sqrt{N}}\int_0^{H_N(S)} F\left(\frac{\widetilde{U}_4^N(s)}{\sqrt{N}}\right)\mathrm{d}s\right).$$

The first term of the right-hand side of Relation (3.58) converges clearly in distribution to 0. The second term can be written as $J_N(S)+M_N(S)$, where, for $t\geq 0$,

$$\begin{aligned}J_N(t)=\kappa_4\int_0^t \mathbb{1}_{\left\{Z_3^N(s=1\right\}} \\ \times\mathbb{E}\left(\int_0^{E_{\kappa_0}} F\left(\frac{Z_4^N(s)-S(Z_4^N(s),u,(E_{\kappa_5,i}))}{\sqrt{N}}\right)\mathrm{d}u\,\middle|\,\mathcal{F}_s\right)\frac{Z_4^N(s)}{\sqrt{N}}\,\mathrm{d}s.\end{aligned}$$

It is easily checked that $(M_N(t))$ is a martingale whose previsible increasing process is given by, for $t\geq 0$,

$$\begin{aligned}\langle M\rangle_N(t) = \frac{\kappa_4}{\sqrt{N}}\int_0^t \mathbb{1}_{\left\{Z_3^N(s)=1\right\}} \\ \times\left(\mathbb{E}\left(\int_0^{E_{\kappa_0}} F\left(\frac{Z_4^N(s)-S(Z_4^N(s),u,(E_{\kappa_5,i}))}{\sqrt{N}}\right)^2\mathrm{d}u\,\middle|\,\mathcal{F}_s\right)\right)\frac{Z_4^N(s)}{\sqrt{N}}\,\mathrm{d}s.\end{aligned}$$

Using Corollary (3.21) one can show that there exists some finite constant C_0 such that

$$\mathbb{E}\left(\langle M\rangle_N(t)\right)\leq\frac{C_0}{\sqrt{N}},$$

the martingale $(M_N(t))$ is converging in distribution to 0.

We are now investigating the asymptotic behavior of $(J_N(t))$.

By using again Corollary 3.21, for $T>0$ and $\varepsilon>0$, there exist constants $0<d_0\leq D_0$ such that

$$\limsup_{N\to+\infty}\mathbb{E}\left(\int_0^T\frac{Z_4^N(s)}{\sqrt{N}}\mathbb{1}_{\left\{Z_4^N(s)\notin[d_0\sqrt{N},D_0\sqrt{N}]\right\}}\,\mathrm{d}s\right)\leq\varepsilon,$$

this is due to the fact that the limit of the occupation measure of $(Z_4^N(s)/\sqrt{N})$ is expressed with a distribution Γ_0 of Definition 3.9 and, in particular, without a mass at 0.

Let $z\geq1$, then

$$\mathbb{E}\left(\int_0^{E_{\kappa_0}}F\left(\frac{z-S(z,u,(E_{\kappa_5,i}))}{\sqrt{N}}\right)\mathrm{d}u\right)=\int_0^{+\infty}\mathbb{E}\left(F\left(\frac{z}{\sqrt{N}}\frac{1}{z}\sum_{i=1}^{z}\mathbb{1}_{\left\{E_{\kappa_5,i}>u\right\}}\right)\right)\mathbb{P}(E_{\kappa_0}\geq u)\,\mathrm{d}u,\tag{3.59}$$

for $\eta>0$ and $u\geq0$, the relation

$$\mathbb{P}\left(\left|\frac{1}{z}\sum_{i=1}^{z}\mathbb{1}_{\left\{E_{\kappa_5,i}>u\right\}}-e^{-\kappa_5u}\right|\geq\eta\right)\leq\frac{1}{z\eta^2}$$

holds. With the uniform continuity of F, we therefore obtain the relation

$$\lim_{N\to+\infty}\sup_{z\in[d_0\sqrt{N},D_0\sqrt{N}]}\mathbb{E}\left(\left|F\left(\frac{z}{\sqrt{N}}\frac{1}{z}\sum_{i=1}^{z}\mathbb{1}_{\left\{E_{\kappa_5,i}>u\right\}}\right)-F\left(\frac{z}{\sqrt{N}}e^{-\kappa_5u}\right)\right|\right)=0.$$

We have therefore that, for the convergence in distribution, the sequence $(J_N(t))$ has the same asymptotic behavior as

$$\left(\kappa_4\int_0^t\mathbb{1}_{\left\{Z_3^N(s-)=1\right\}}\int_0^{+\infty}F\left(\frac{Z_4^N(s)}{\sqrt{N}}e^{-\kappa_5u}\right)e^{-\kappa_0u}\,\mathrm{d}u\frac{Z_4^N(s)}{\sqrt{N}}\,\mathrm{d}s\right).$$

We can now use Corollary (3.21) and standard calculus to complete the proof of the theorem. □

We now establish the fact that the process $(X_2^N(t))$ has indeed the same asymptotic behavior as $(U_2^N(t))$.

PROPOSITION 3.27. *Let $(X_N(t))$ and $(U_N(t))$ be the solutions of the SDEs* (3.40) *and* (3.51) *with initial point $(0,y_N,1,v_N)$ and (y_N,v_N) satisfies Relation* (3.41)*, then the two sequences*

$$\left(\frac{X_2^N(\sqrt{N}t)}{N}\right)\ \textit{and}\ \left(\frac{U_2^N(\sqrt{N}t)}{N}\right)$$

have the same limit for the convergence in distribution.

PROOF. The proof follows the arguments used for the convergence in distribution of $(U_2^N(\sqrt{N}t)/N)$. Some adjustments are nevertheless necessary but the main ideas are essentially the same. We sketch the main lines of the proof.

If $(X_N(t))$ is a solution of SDEs (3.40), on the time interval $[0,\sqrt{N}T]$ the contribution of the Poisson process $\mathcal{P}_1$ to the coordinate $(X_2^N(t))$ is of the order of $\sqrt{N}$ which is negligible since the order of magnitude considered for $(X_2^N(t))$ is N. Therefore, we can take this Poisson process out of the set of SDEs for $(X_N(t))$.

For $T>0$, it is not difficult to show that, with high probability, the values of the process $(X_2(\sqrt{N}t)/N)$ are in $(\delta,3/2)$ on the time interval $[0,T]$. When there is a new arrival for the chemical species S_1, it is transformed into chemical species S_3

at rate at least $\kappa_2 \delta N$. Hence, with high probability, on the time interval $[0, \sqrt{N}T]$, the values of the process $(X_1^N(t))$ are in the set $\{0, 1\}$.

A central argument for the convergence of $(U_2^N(\sqrt{N}t)/N)$ is Theorem 3.20. We define by $(\widetilde{Z}_N(t))$ the analogue of $(Z_N(t))$, i.e. the time-changed process $(X_N(t))$ with all time intervals where X_3^N is null are removed. The process $(\widetilde{Z}_N(t))$ satisfies the analogue of the SDEs (3.52) where the last SDE is replaced by

$$\begin{aligned}\mathrm{d}\widetilde{Z}_4^N(t) = \mathcal{P}_3\left((0, \kappa_3 \widetilde{Z}_2^N \widetilde{Z}_3^N(t-)), \mathrm{d}t\right) - \mathcal{P}_5\left((0, \kappa_5 \widetilde{Z}_4^N(t-)), \mathrm{d}t\right),\\ -\mathbb{1}_{\left\{\widetilde{Z}_3^N(t-)=1\right\}} \int_{a,b} S\left(\widetilde{Z}_4^N(t-), a + \frac{c}{\kappa_2 \widetilde{Z}_2^N(t-)}, b\right)\\ \times \widetilde{\mathcal{N}}\left(\left(0, \kappa_4 Z_4^N(t-)\right], \mathrm{d}t, \mathrm{d}a, \mathrm{d}b, \mathrm{d}c\right),\end{aligned}$$

where $\widetilde{\mathcal{N}}$ be a Poisson marked point process on $\mathbb{R}_+^2 \times \mathbb{R}_+ \times \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+ \times$ with intensity measure

$$\mathrm{d}s \otimes \mathrm{d}t \otimes \kappa_0 e^{-\kappa_0 a}\, \mathrm{d}a \otimes Q(\mathrm{d}b) \otimes e^{-c}\, \mathrm{d}c,$$

where, as before, Q is the distribution of (E_i) on $\mathbb{R}_+^{\mathbb{N}}$. The additional variable c of the Poisson process $\widetilde{\mathcal{N}}$ in this SDE is due to the fact that when $\widetilde{Z}_1^N(t)=1$ and $\widetilde{Z}_3^N(t)=0$, $\widetilde{Z}_1$ leaves the state 1 at rate $\kappa_2 \widetilde{Z}_2^N(t-)$. Because of the assumption on $(\widetilde{Z}_2^N(t-/N)$ in $(\delta, 3/2)$, a glance at the proof of Theorem 3.12 shows that, even with this extra term, the limit result of this theorem still holds with the same limits. The proof is then concluded as in the proof of Proposition 3.23. □

CHAPTER 4

Analysis of Stochastic Chemical Reaction Networks with a Hierarchy of Timescales

Contents

1. Introduction

A *stochastic chemical reaction network* (CRN) with n chemical species is described as a continuous time Markov process $(X_i(t))$ on a subset of $\mathbb{N}^n$. The ith component gives the number of molecules of chemical species S_i, $1 \le i \le n$. Its dynamical behavior is given by a finite set of chemical reactions which add or remove simultaneously a finite number of several chemical species. For example, the reaction

$$k_1 S_1 + k_2 S_2 \xrightarrow{\kappa} k_3 S_3 \tag{4.1}$$

transforms k_1 molecules of S_1 and k_2 molecules of S_2 into k_3 molecules of S_3. The associated transition of this reaction for the Markov process is

$$x = (x_i) \to x + k_3 e_3 - k_1 e_1 - k_2 e_2,$$

where e_i, $1 \le i \le n$, is the ith unit vector of $\mathbb{N}^n$. The rate at which the reaction occurs is assumed to follow the *law of mass action*, for our example the rate is given by

$$\kappa x_1^{(x_1)} x_2^{(x_2)} \stackrel{\text{def.}}{=} \kappa \frac{x_1!}{(x_1 - k_1)!} \frac{x_2!}{(x_2 - k_2)!}, \tag{4.2}$$

for some positive constant κ. See Section 2.

From a mathematical point of view, there are two important characteristics of stochastic models of CRNs described with Markov processes.

(a) Polynomial Reaction Rates.
When the coordinates x_1 and x_2 are large, the reaction rate (4.2) is of the order of $\kappa {x_1}^{k_1} {x_2}^{k_2}$. This implies that some reactions will be much more likely than others, and therefore will dominate the kinetics of the CRN, for a while at least. In this case, we will speak of fast processes for the coordinates involved in these reactions. There are many examples of such behavior. See Agazzi and Mattingly [4], Ball et al. [11], Togashi and Kaneko [74] and Sections 6, 7, 8 of Laurence and Robert [52] for example. This is a major feature of CRNs from a technical point of view. In such a case, a CRN can be described as driven by a set of interacting fast processes

leading to an investigation of possible stochastic averaging principles or even more complex multi-timescales behaviors. See Section 1.3.

(b) Boundary Behavior.
This feature is due to a constraint on the state space rather than a property related to the order of magnitude of transition rates. In state $x=(x_i)\in\mathbb{N}^n$, Reaction (4.1) occurs only if $x_1\geq k_1$ and $x_2\geq k_2$. Mathematically, this is a kind of discontinuity of the kinetics of the CRN. This constraint on the state space is at the origin of complex behaviors of CRNs. In the CRN of example (4.1), if we assume that $X_1(0)=N$ is large and that the process $(X_2(t))$ remains in a neighborhood of 0, then the process $(X_1(t))$ will decrease only during the excursions of $(X_2(t))$ above k_2. This can be even more complicated if the dynamic of $(X_2(t))$ depends, via other chemical reactions, on $(X_3(t))$ for example. For example of such complex behaviors, see Section 8 of [52] and Laurence and Robert [54].

1.1. k-Unary Chemical Reaction Networks. We now describe the class of CRNs analyzed in our paper. As it will be seen boundary behaviors play only a marginal role in the time evolution of these networks. The characteristic (a) on the polynomial growth is the key feature.

The parameters of the kinetics of these networks are given by the coefficients of a matrix $R_\kappa=(\kappa_{ij}, 0\leq i,j\leq n)\in\mathbb{R}_+^{n+1}\times\mathbb{R}_+^{n+1}$, and a vector $(k_i)\in(\mathbb{N}\setminus\{0\})^n$ of integers. The only chemical reactions for this class of CRNs are as follows, for $1\leq i\neq j\leq n$,

$$k_iS_i \xrightarrow{\kappa_{ij}} k_jS_j, \quad k_iS_i \xrightarrow{\kappa_{i0}} \emptyset, \quad \emptyset \xrightarrow{\kappa_{0i}} k_iS_i,$$

provided that, respectively, $\kappa_{ij}>0$, $\kappa_{i0}>0$, or $\kappa_{0i}>0$. The second reaction, resp. last reaction, is the spontaneous destruction, resp. creation, of k_i molecules of chemical species S_i. The symbol $\emptyset$ is the source/sink for chemical species.

For $1\leq i\leq n$, k_iS_i is the only *complex* involving the chemical species S_i and the time evolution of the ith coordinate is a jump process whose jumps are $\pm k_i$. In state $x=(x_k)$, for $i\in\{1,\ldots,n\}$, the ith coordinate decreases at a rate proportional to $x_i^{(k_i)}$ and, for $1\leq j\leq n$, $\kappa_{ij}x_i^{(k_i)}$ is the rate at which k_i molecules of S_i are transformed into k_j molecules of S_j. These are the kinetics of the law of mass action. See Section 2.

This class of CRNs has in fact an invariant distribution, see Relation (4.13) of Section 2.5, given by a product of Poisson distributions. If this is satisfactory, it should be noted that there are many very different Markov processes with this property, see [61]. It does not give much insight on the transient characteristics of the CRNs, in particular on the impact of its different timescales of this CRNs, if any.

A scaling approach is proposed to investigate the dynamical behavior of these networks. We quickly review several scalings already used in the literature of stochastic CRNs.

1.2. Scaling Methods for Chemical Reaction Networks. We denote by N the scaling parameter.

(a) Classical Scaling.
For this scaling the reaction rate κ_r of a chemical reaction r, is scaled in N, as κ_r/N_r^γ for some $\gamma_r\geq 0$, so that if all coordinates of the associated Markov process $(X(t))=(X_i(t))$ are of the order of N, then the transition rate of any jump of the process is of the order of N. See Mozgunov et al. [60] or Proposition 2 of Laurence and Robert [52] for example. In this case, under appropriate conditions, it can be shown that the process $(X_i^N(t)/N)$ is converging in distribution to the solution of an ODE whose

stability properties have been investigated in the literature of deterministic CRNs. See Feinberg [26] and Horn and Jackson [42] for example.

This scaling has the effect of somewhat equalizing the kinetics of the CRNs. There cannot be a subset of chemical reactions dominating at some moment for a while, since all transition rates are of the order of N.

Kurtz and co-authors have also investigated several examples of CRNs with related scaling methods. In this approach, some reaction rates may be sped-up with some power of the scaling parameter and the state variables are scaled accordingly. There is no requirement that all reactions have the same order of magnitude. The initial motivation was of fitting the parameters of these scaling models with biological data obtained from experiments. See for example Ball et al. [11], Kang and Kurtz [46], and Kim et al. [50] where, for several examples of CRNs, the choice of convenient scalings of reaction rates is investigated and several limit theorems are derived.

(b) Scaling with the norm of the initial state.
In this approach the reaction rates κ_r are fixed so that the topology of the CRN is preserved by the scaling. The scaling parameter for the Markov process $(X(t))$ is $N{=}\|X(0)\|$. The approach consists in describing, via possibly functional limit theorems, how the sample path of the state of the CRN returns to a neighborhood of the origin. This is a natural way to investigate positive recurrence properties of the CRNs but, more importantly, it can provide insight into transient characteristics of CRNs. Up to now there are few results in the literature in this domain, see Agazzi et al. [1] and [3], Mielke et al. [59], and McSweeney and Popovic [57]. For the scaling with $\|X(0)\|$, see Laurence and Robert [52] and references therein.

Scaling External Input Rates. The scaling investigated in this paper is as follows. For all $i{\in}\{1,\dots,n\}$ such that $\kappa_{0i}{>}0$, the creation of chemical species S_i is scaled by N, it becomes

$$\emptyset \xrightarrow{N\kappa_{0i}} k_i S_i.$$

The other reaction rates do not change. Rather than starting from a "large" initial state, this scaling regime assume heavy traffic conditions at the entrance of the CRNs. A natural question in this setting is of establishing a limit theorem on the orders of magnitude in N of the coordinates of $(X_N(t)){=}(X_i^N(t))$. This scaling has already been considered in Togashi and Kaneko [73] for CRNs and in Ball et al. [11], and probably in many other examples. A related scaling has also been used to investigate the transient behavior of Markov processes for stochastic models of large communication networks in Kelly [62]. See also Kelly [61] for a survey.

A basic example of such a situation is the k-unary CRN with one chemical species,

$$\emptyset \underset{\mu}{\overset{\lambda N}{\rightleftharpoons}} k_1 S_1.$$

It can be easily seen that, under convenient initial conditions, the scaled process

$$\left(\frac{X_1^N\left(t/N^{1-1/k_1}\right)}{N^{1/k_1}} \right) \tag{4.3}$$

converges in distribution to a non-trivial deterministic function, the solution of an ODE. See Proposition 4.5.

A Hierarchy of Timescales. We come back to our CRNs under the heavy traffic assumptions, i.e. with all external input rates scaled by N. Heuristically, if there is a kind of equilibrium of flows in the network at some moment, due to the external inputs of the order of N, the input flow through each node should be also of the same order of N.

The case of the CRN with a single node suggests then that the state variable of the ith node $(X_i^N(t))$, $1 \leq i \leq n$, should be of the order of N^{1/k_i}. The convergence result for the process (4.3) indicates that the "natural" timescale of $(X_i(t))$ should be $(t/N^{1-1/k_i})$. In particular, this implies that, at the "normal" timescale (t), all coordinates $(X_i(t))$ whose index $i \in \{1, \ldots, n\}$ is such that $k_i \geq 2$, are fast processes. The CRN exhibits in fact a hierarchy of timescales: The process associated to $(X_j(t))$ is faster than the process $(X_i(t))$ provided that $k_j > k_i$. A limit theorem to establish the convergence of the scaled process

$$\left(\frac{X_i^N(t)}{N^{1/k_i}} \right) \tag{4.4}$$

has to handle this multi-timescales feature and also the interactions with the other coordinates.

1.3. Literature. A classical way of investigating multi-timescales processes is via the proof of an averaging principles. Averaging principles have already been used in various situations to study chemical reaction networks (CRNs). In most of cases, it involves two timescales: there are a fast process and a slow process. Early works on the proof of averaging principles are due to Has'minskiĭ. See Has'minskiĭ [39, 40]. Chapter 7 of Freidlin and Wentzell [31] considers these questions in terms of the convergence of Cesaro averages of the fast component. Papanicolaou et al. [63] has introduced a stochastic calculus approach to these problems, mainly for diffusion processes. Kurtz [55] has extended this approach to jump processes. For CRNs, there are numerous proofs of averaging principles in such a setting: Ball et al. [11], Kang and Kurtz [46], Kim et al. [50], Laurence and Robert [52, 54], ...

With more than two timescales, limit theorems in a stochastic framework are quite scarce in the literature. A model with three timescales is investigated in Kang et al. [47], and a functional central limit result is established. In this reference, it is assumed that the first order is deterministic. To handle the two fast timescales, several assumptions on uniform convergence of infinitesimal generators on compact subsets of the state space are introduced. Large deviations results are derived with similar assumptions in Popovic [65]. It does not seem that such an approach can be used in our case.

A stochastic model of a CRN with three timescales is analyzed in Fromion et al. [32]. The limiting behavior of the occupation measure of the processes associated to the two fast timescales is investigated. The main difficulty is of identifying the possible limits. A technical result on conditional probabilities is the major ingredient to solve this problem. This method do not seem to be possible for our CRN, mainly because there are too many fast timescales a priori, so that an analogous result on conditional probabilities is not clear.

1.4. Outline of the Paper. The goal of this paper is of establishing a limit theorem for the convergence in distribution of the scaled process defined by Relation (4.4) :

- For the occupation measure of the coordinates of the Markov process whose indices $i \in \{1, \ldots, n\}$ are such that $k_i \geq 2$;
- For the vector of the other components, i.e. indices $i \in \{1, \ldots, n\}$ with $k_i = 1$, for the uniform topology.

See Theorem 4.4 for the full statement. The proof of this result is done in several steps.

(a) Technical estimates of the "basic" model of a k-unary CRN with one chemical species in Section 3;
(b) Tightness results for the occupation measure by using (a) and linear algebra arguments in Section 4;
(c) Identification of the limit of the sequence of occupation measures. This is done first by establishing a functional equation for some marginals of the possible limiting points, Relation (4.40) of Proposition 4.14, and then by induction on the hierarchy of timescales starting from the fastest timescale. Relative entropy functions associated to each timescale and convexity arguments are the main ingredients of the proofs. In Section 5 when all k_i, $i=1,\dots,n$, are greater than 2, and Section 6 for the general case.

2. Stochastic Model

We introduce the formal definitions and notations used throughout the paper.

2.1. The class of k-unary chemical reaction networks.

Definition 4.1 (k-unary CRN). *The components of a* k-unary chemical reaction network *are :*

(a) *A set of* n *distinct* chemical species $\mathcal{S}=\{S_1,\dots,S_n\}$. *The set* $\mathcal{S}$ *is also identified to* $\{1,\dots,n\}$ *and* $\emptyset$ *is the source/sink for chemical species, it is associated to index* $i=0$ *in general;*
(b) Complexes $\mathcal{C}$ *are of the form* k_iS_i, $i=1,\dots,n$, *for some* $k_i\geq 1$. *We will have the convention* $k_0=0$. *Each species is present in exactly one complex.*
(c) *The* rates of chemical reactions *are associated to a* Q*-matrix* $R_\kappa=(\kappa_{ij},i,j\in I)$ *of a jump Markov process on* $I=\{0,\dots,n\}$ *in the following way: If* i, $j\in I$ *are such that* $\kappa_{ij}>0$, *then there is the reaction*

$$\begin{cases} k_iS_i \xrightarrow{\kappa_{ij}} k_jS_j & \text{if } i\neq 0;\\ \emptyset \xrightarrow{\kappa_{0j}N} k_jS_j & \text{if } i=0, \end{cases}$$

where N *is the* scaling parameter. *These are the only possible reactions.*

Note that the process associated to the Q-matrix R_κ *is not* the process describing the time evolution of the CRN, it is a jump process on the finite set I. The state of the CRN is given by $(X_N(t))=(X_i^N(t))$, a Markov process with values in $\mathbb{N}^n$. Since, for $i\in\{1,\dots,n\}$, the sizes of jumps of the number of copies of chemical species i are either $\pm k_i$, a natural state space for this process is

$$\mathcal{S}_a=\{x=(x_i)=(a_1+m_1k_1,a_2+m_2k_2,\dots,a_n+m_nk_n):(m_i)\in\mathbb{N}^n\}, \tag{4.5}$$

for any $a\in\{0,\dots,k_1-1\}\times\{0,\dots,k_2-1\}\times\cdots\times\{0,\dots,k_n-1\}$.

The kinetics of the system are driven by *the law of mass action*, see Voit et al. [76], Lund [56] for surveys on the law of mass action and the historical reference Guldberg and Waage [35]. The associated transitions are thus given by, for $x\in\mathcal{S}_a$, i, $j\in I$, $i\neq 0$,

$$x=(x_i)\to x+\begin{cases} k_je_j-k_ie_i, & \text{at rate} \quad \kappa_{ij}x_i^{(k_i)}\\ k_ie_i, & \text{“} \quad \kappa_{0i}N\\ -k_ie_i, & \text{“} \quad \kappa_{i0}x_i^{(k_i)}. \end{cases}$$

where e_i is the ith unit vector of $\mathbb{N}^n$ and, for y, $k\in\mathbb{N}$,

$$y^{(k)}=\frac{y!}{(y-k)!}, \tag{4.6}$$

if $y \geq k$ and $y^{(k)}=0$ otherwise.

Such CRNs have a *fast input*, in the sense that the rates of creations of chemical species are proportional to a (large) scaling factor N, and these are the only chemical reactions which are sped-up.

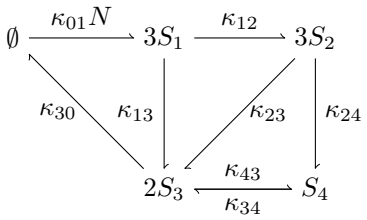


FIGURE 1. An example of a k-unary CRN

2.2. Relations with Some Queueing Networks. A k-unary network can be related to several queueing systems.

(a) When there is one chemical species, this is a generalized $M/M/\infty$ queue. See Section 3.

(b) Queueing networks referred to as *Jackson Networks*. They can be described simply as follows.
   — There are n sites for the location of jobs. If $x=(x_j)\in\mathbb{N}^n$, for $1\leq j\leq n$, x_j denotes the number of jobs at the site j.
   — One of the jobs at site i leaves at rate $\mu_{ij}>0$ to go to site $j\in\{1,\ldots,n\}$, or leave the network at rate μ_{i0}
   — External jobs arrive at the site i at rate $\mu_{0i}\geq 0$.

   The main difference with our CRNs is that the ith coordinate, $i\in\{1,\ldots,n\}$ decreases at a fixed rate if it is not 0, instead of a rate proportional to $x_i^{(k_i)}$ for a k-unary CRN. There is a scaling result with the norm of the initial state for these networks in Chen and Mandelbaum [66], see Section 1.2. The scaling results are quite different, there is only one timescale for Jackson networks. Nevertheless, as for our CRNs, a linear system plays an important role in the limit theorems associated to these Markov processes. See Relation (4.28) in Proposition 4.9 for k-unary CRNs and, for Jackson networks, see Proposition 9.6 of Robert [67].

2.3. Notations. Throughout the paper, the following notations will be used. For a subset A of $\mathbb{R}$, we denote $A^*=A\backslash\{0\}$ and, for p, $q\in\mathbb{N}^*$, $p\leq q$,

$$I_{[p,q]} \stackrel{\text{def.}}{=} \{0\}\cup\{i\geq 1 : p\leq k_i\leq q\}, \tag{4.7}$$

with the convention that $I_{[p]}=I_{[p,p]}$, and $I_{[p+]}=I_{[p,+\infty]}$, so that $I=I_{[1+]}$.

If $x\in\mathbb{R}_+^{I^*}$ and $1\leq p\leq q$, we define $x_{[p,q]}=(x_i, i\in I^*_{[p,q]})$ and x will also be represented as $x=(x_{[k_1]}, x_{[k_2]},\ldots, x_{[k_n]})$ or $x=(x_{[1,q-1]}, x_{[q+]})$, for $q\geq 2$, provided that $I^*_{[1,q-1]}$ and $I^*_{[q+]}$ are non-empty. Similarly, if π is a probability distribution on on $\mathbb{R}_+^{I^*}$, $\pi^{[p,q]}$ is the distribution on $\mathbb{R}_+^{I^*_{[p,q]}}$ of marginals of π for the coordinates whose index is in $I^*_{[p,q]}$, i.e. the image of π by the mapping $x\mapsto x_{[p,q]}$.

For any subset A of I, we denote by $\Omega(A)$ the set of irreducible Q-matrices $(x_{ij}, i,j\in A)$ for the state space A. We will assume in this paper that $R_\kappa\in\Omega(I)$, with a slight abuse of notation we will also write $\kappa\in\Omega(I)$. An $A\times A$-matrix refers

to a $|A|\times|A|$-matrix, with $|A|$ the cardinality of A. For $i\in I^*$, we define

$$\kappa_i^+=\kappa_{i0}+\sum_{j\in I^*\setminus\{i\}}\kappa_{ij}. \tag{4.8}$$

We now introduce a natural distance $(d(i))$ from the origin (the complex $\emptyset$) on the graph of the CRN.

DEFINITION 4.2. *We set* $d(0)=0$ *and, for* $1\le i\le n$,

$$d(i)=\min\left\{k\ge 1:\exists i_1,\dots,i_{k-1}\in I^*,\kappa_{0,i_1}\cdot\prod_{p=1}^{k-2}\kappa_{i_pi_{p+1}}\cdot\kappa_{i_{k-1}i}>0\right\}.$$

A real-valued function $(x(t))$ on $(\mathbb{R}_+)^{I^*}$ is càdlàg if it is right continuous and it has left-limits everywhere on $\mathbb{R}_+^*$, in this case, for $t>0$, $x(t-)$ denotes the left limit of $(x(t))$ at $t>0$. If H is a subset of $\mathbb{R}^d$, for $d\ge 1$, we denote by $\mathcal{B}(H)$ the set of Borelian subset of H, $\mathcal{C}_c(H)$ the set of continuous functions on H with compact support on H and $\mathcal{C}_c^2(H)$ the subset of class $\mathcal{C}_2$-functions and the set on Borelian probability distributions on H is denoted as $\mathcal{P}(H)$.

The convergence in distribution of a sequence of jump processes $(U_N(t))$ in $\mathbb{R}^d$ to a process $(U(t))$ is understood with respect to the topology of uniform convergence on compact sets for càdlàg functions. See Chapters 2 and 3 of Billingsley [14] for example. The convergence in distribution of the associated occupation measures is the convergence in distribution of the sequence of random measures (μ_N) on $\mathbb{R}_+^d$, defined by, for $f\in\mathcal{C}_c(\mathbb{R}_+^{d+1})$,

$$\langle\mu_N,f\rangle=\int_0^T f(s,U_N(s))\,\mathrm{d}s.$$

See Dawson [21] for the technical aspects related to measure valued processes.

2.4. Stochastic Differential Equations. If $\mathcal{P}$ is a positive Borelian measure on $\mathbb{R}_+^2$, and $A\in\mathcal{B}(\mathbb{R}_+)$ is a Borelian subset of $\mathbb{R}_+$, we use the following notation,

$$\mathcal{P}(A,\mathrm{d}t)=\int_{x\in\mathbb{R}_+}\mathbb{1}_{\{x\in A\}}\mathcal{P}(\mathrm{d}x,\mathrm{d}t). \tag{4.9}$$

We will express the time evolution of the k-unary CRN, as a càdlàg process $(X_N(t))=(X_i^N(t),i=1,\dots,n)$, solution of the following stochastic differential equation (SDE). See Laurence and Robert [52]. For $i\in I^*$, $t\ge 0$,

$$\begin{aligned}\mathrm{d}X_i^N(t)=k_i\mathcal{P}_{0i}\left((0,\kappa_{0i}N),\mathrm{d}t\right)&+\sum_{j\in I^*\setminus\{i\}}k_i\mathcal{P}_{ji}\left(\left(0,\kappa_{ji}(X_j^N(t-))^{(k_j)}\right),\mathrm{d}t\right)\\&-\sum_{j\in I\setminus\{i\}}k_i\mathcal{P}_{ij}\left(\left(0,\kappa_{ij}(X_i^N(t-))^{(k_i)}\right),\mathrm{d}t\right).\end{aligned} \tag{4.10}$$

where $\mathcal{P}_{ij}$, $i,j\in I$ is a family of independent Poisson point processes on $\mathbb{R}_+^2$ with intensity measure the Lebesgue measure on $\mathbb{R}_+^2$. See Kingman [64].

The martingale and stopping time properties will refer to the smallest filtration $(\mathcal{F}_t)$ satisfying the usual hypotheses and such that

$$\{\mathcal{P}_{ij}(A\times[0,s]):i\in I,j\in I\setminus\{i\},A\in\mathcal{B}(\mathbb{R}_+),s\le t\}\subset\mathcal{F}_t,\quad\forall t\ge 0.$$

2.5. Invariant Distribution with Product Form Representation. In the language of chemical reaction networks, a k-unary CRN is *weakly reversible* with *one linkage class* and its *deficiency* is 0. See Feinberg [27] for the general definitions for CRNs.

The Deterministic CRN.
In a deterministic setting, a dynamical system $(u_N(t))=(u_i^N(t))$ on $\mathbb{R}_+^n$ is associated to this CRN

$$\frac{\dot{u}_i^N(t)}{k_i} = N\kappa_{0i} + \sum_{j\in I^*\setminus\{i\}} \left(u_j^N(t)\right)^{k_j} \kappa_{ji} - (u_i^N(t))^{k_i} \sum_{j\in I\setminus\{i\}} \kappa_{ij}, \quad i\in I^*. \tag{4.11}$$

Classical results of Feinberg [26] and Horn and Jackson [42] show that, in this case, $(u_N(t))$ has a unique equilibrium point $\gamma_N=(N^{1/k_i}u_i)$ which is locally stable, where $u_\infty=(u_i)$ is the unique positive solution of the system of equations,

$$\kappa_{0i} + \sum_{j\in I^*\setminus\{i\}} u_j^{k_j}\kappa_{ji} = u_i^{k_i} \sum_{j\in I\setminus\{i\}} \kappa_{ij}, \quad i\in I^*. \tag{4.12}$$

See Proposition 4.9 of Section 4.1 and Feinberg [27] for a general presentation of these dynamical systems.

The invariant Measure.
For $a=(a_i)\in\mathbb{N}^n$, with $a_i\in\{0,\dots,k_i-1\}$ for all $1\le i\le n$, the Markov process $(X_N(t))$ is irreducible on the set $\mathcal{S}_a$ defined by Relation (4.5). Anderson et al. [7] shows that the invariant distribution of $(X_N(t))$ on $\mathcal{S}_a$ is given by

$$\nu_a(x) = \frac{1}{Z_a}\prod_{i=1}^n \frac{(\gamma_{j,N})^{x_i}}{x_i!}, x\in\mathcal{S}_a, \tag{4.13}$$

where Z_a is the normalization constant,

$$Z_a = \sum_{k=(k_i)\in\mathbb{N}^n} \prod_{i=1}^n \frac{(\gamma_{i,N})^{a_i+p_ik_i}}{(a_i+p_ik_i)!}.$$

and $\gamma_N=(\gamma_{i,N})=(N^{1/k_i}u_i)$, where (u_i) is the solution of the system (4.12).

2.6. Timescales. When $n=1$, the k-unary CRN is

$$\emptyset \underset{\kappa_{10}}{\overset{\kappa_{01}N}{\rightleftharpoons}} kS_1,$$

in state x, the instantaneous mean drift of X_N is $k(\kappa_{01}N-\kappa_{10}x^{(k)})$. In view of Relation (4.6), to have a non-trivial time evolution when N is large, this suggests that x should be of the order of $N^{1/k}$. It is not difficult to show that, provided that the sequence $(X_1^N(0)/N^{1/k})$ converges, then the sequence of processes

$$\left(\frac{X_1^N(t/N^{1-1/k})}{N^{1/k}}\right)$$

is converging in distribution to $(x_1(t))$ the solution of the ODE

$$\dot{x}_1(t) = k(\kappa_{01}-\kappa_{10}x_1(t)^k), \quad t\ge 0.$$

See Section 3. The natural timescale of the process $(X_1^N(t)/N^{1/k})$ is $(t/N^{1-1/k})$. If $k\ge 2$, $(X_1^N(t)/N^{1/k})$ is then a *fast process*, and when $k=1$, $(X_1^N(t)/N)$ can be seen as a *slow process*.

For our general k-unary CRN, fast and slow processes define a partition of the set of indices $i\in\{1,\dots,n\}$ based on the fact that $k_i=1$ or $k_i\ge 2$, i.e. $I^*=I^*_{[1]}\cup I^*_{[2+]}$. In the same way, if i, $j\in I^*$, is such that $k_i>k_j$, then the process $(X_i^N(t)/N^{1/k_i})$ is "faster" than the process $(X_j^N(t)/N^{1/k_j})$. This leads to a classification of chemical species according to their natural timescales, i.e. according to the value of k_i. This hierarchy plays an important role in the proofs of convergence in distribution of this paper.

2.7. The Convergence Result. With the above remark, the set I^*_{2+} is the set of indices of fast processes, the asymptotic evolution of $(X^N_i(t), i{\in}I^*_{[2+]})$ is described only in terms of its *occupation measure*. For I^*_1, the set of indices associated to slow processes, this is the convergence in distribution of the sequence of processes $(\overline{X}^N_i(t), i{\in}I^*_{[1]})$.

Definition 4.3.

(a) *The scaled process* $(\overline{X}^N(t))$*, is defined for* $N{\geq}1$ *as*

$$\left(\overline{X}_N(t)\right) = \left(\overline{X}^N_i(t)\right) = \left(\frac{X^N_i(t)}{N^{1/k_i}}\right). \tag{4.14}$$

The initial state $X_N(0){=}x_N{=}(x^N_i){\in}\mathbb{N}^n$ *of the process* $(X_N(t))$ *is assumed to satisfy the relation*

$$\lim_{N\to+\infty}\left(\frac{x^N_i}{N^{1/k_i}}\right) = (\alpha_i) \in \left(\mathbb{R}^*_+\right)^n. \tag{4.15}$$

(b) *The* occupation measure Λ_N *is the random measure on* $\mathbb{R}_+{\times}\mathbb{R}^{I^*}_+$ *defined by, for* $g{\in}\mathcal{C}_c(\mathbb{R}_+{\times}(\mathbb{R}^*_+)^{I^*})$*,*

$$\langle \Lambda_N, g\rangle = \int_{\mathbb{R}_+} g\left(u, \left(\overline{X}^N_i(u), i{\in}I^*\right)\right) \mathrm{d}u. \tag{4.16}$$

The main result of the paper is the following theorem.

Theorem 4.4. *If* $(X_N(t))$ *is the solution of SDE* (4.10) *whose initial condition satisfies Condition* (4.15)*, then, for the convergence in distribution,*

$$\lim_{N\to+\infty}\left(\left(\overline{X}^N_i(t), i{\in}I^*_{[1]}\right), \Lambda_N\right) = \left(\left(x_i(t), i{\in}I^*_{[1]}\right), \Lambda_\infty\right), \tag{4.17}$$

where $(\overline{X}_N(t))$ *and occupation measure* Λ_N *are defined respectively by Relations* (4.14) *and* (4.16)*, with, for* $g \in \mathcal{C}_c\left(\mathbb{R}_+{\times}(\mathbb{R}^*_+)^{I^*}\right)$*,*

$$\langle \Lambda_\infty, g\rangle = \int_{\mathbb{R}_+} g\left(s, \left(x(s), \left(L_i(x(s)), i{\in}I^*_{[2+]}\right)\right)\right) \mathrm{d}s, \tag{4.18}$$

where:

(a) *If* $y{\in}(\mathbb{R}^*_+)^{I^*_{[1]}}$*,* $L(y){=}(L_i(y), i{\in}I^*_{[2+]})$ *is the unique solution of the system*

$$\kappa_{0i} + \sum_{j\in I^*_{[1]}} y_j\kappa_{ji} + \sum_{j\in I^*_{[2+]}\setminus\{i\}} L_j(y)^{k_j}\kappa_{ji} = L_i(y)^{k_i}\sum_{j\in I\setminus\{i\}}\kappa_{ij}, \quad i{\in}I^*_{[2+]}; \tag{4.19}$$

(b) *The function* $(x(t)){=}(x_i(t), i{\in}I^*_{[1]})$ *is the unique solution of the set of ODEs,*

$$\begin{aligned}\dot{x}_i(t) = \kappa_{0i} &+ \sum_{j\in I^*_{[1]}\setminus\{i\}} x_j(t)\kappa_{ji} \\ &+ \sum_{j\in I^*_{[2+]}\setminus\{i\}} L_j(x(t))^{k_j}\kappa_{ji} - x_i(t)\sum_{j\in I\setminus\{i\}}\kappa_{ij}, \quad i{\in}I^*_{[1]},\end{aligned} \tag{4.20}$$

with initial point $(\alpha_i, i{\in}I^*_{[1]})$*.*

Not that the coordinates of the scaled vector $(\overline{X}_N(t))$ with indices in $I_{[1]}$ also appear in Λ_N even if there is a much stronger result for the convergence in distribution for them. This is only to have simpler expressions.

3. A Generalized $M/M/\infty$ Queue

In this section, we will study the simplest form of k-unary CRN, a CRN with only one species,

$$\emptyset \underset{\mu}{\overset{\lambda N}{\rightleftharpoons}} kS.$$

The process $(X_N(t))$ is a birth and death process with the transition rates, for $x \geq 0$,

$$x \longrightarrow x + \begin{cases} k & \text{at rate } \lambda N, \\ -k & \text{“} \qquad \mu x^{(k)}. \end{cases} \tag{4.21}$$

When $k=1$, $(X_N(t))$ is the Markov process of the $M/M/\infty$ queue, with arrival rate λN, and departure rate μ. It is a basic model in the study of stochastic chemical reaction networks. See Laurence and Robert [52] and Chapter 6 of Robert [67] for a general presentation.

We start with a simple scaling result.

Proposition 4.5. *If the initial condition x_n of the Markov process $(X_N(t))$ is such that*

$$\lim_{N \to +\infty} \frac{x_n}{\sqrt[k]{N}} = \alpha,$$

then, for the convergence in distribution, the relation

$$\lim_{N \to +\infty} \left(\frac{1}{\sqrt[k]{N}} X_N \left(t/N^{1-1/k} \right), t \geq 0 \right) = \left(x(t), t \geq 0 \right),$$

holds, where $(x(t))$ is the solution of the ODE $\dot{x}(t) = \lambda - \mu x(t)^k$, with $x(0) = \alpha$.

Proof. This is done with straightforward stochastic calculus. The SDE (4.10) is in this case

$$\mathrm{d}X_N(t) = k\mathcal{P}_{01}((0, \lambda N), \mathrm{d}t) - k\mathcal{P}_{10}((0, \mu X_N(t-)^{(k)}), \mathrm{d}t), \tag{4.22}$$

by integrating this relation, we obtain that, for $t \geq 0$,

$$Y_N(t) \overset{\text{def.}}{=} \frac{1}{\sqrt[k]{N}} X_N \left(t/N^{1-1/k} \right) = Y_N(0) + M_N(t) + \lambda k t - k \int_0^t \frac{X_N(s)^{(k)}}{N} \, \mathrm{d}s, \tag{4.23}$$

where $(M_N(t))$ is a martingale whose previsible increasing process is given by

$$\langle M_N \rangle (t) = \frac{\lambda k^2 t}{N^{1+1/k}} + \frac{\mu k^2}{N^{1/k}} \int_0^t \frac{X_N(s)^{(k)}}{N} \, \mathrm{d}s,$$

therefore, with Relation (4.23) we get

$$\mathbb{E}\left(\langle M_N \rangle (t) \right) \leq \frac{\lambda k^2 t}{N^{1+1/k}} + \frac{\mu k}{N^{1/k}} \left(Y_N(0) + \lambda k t \right).$$

Doob's Inequality gives that the sequence of martingales $(M_N(t))$ is converging in distribution to 0. By using again Relation (4.23), we get that, for any $T > 0$ and $\varepsilon > 0$, there exists K such that

$$\mathbb{P}\left(\sup_{t \leq T} Y_N(t) \geq K \right) \leq \varepsilon.$$

We can then use the criterion of the modulus of continuity, see Theorem 7.3 of Billingsley [14], to show that the sequence $(Y_N(t))$ is tight for the convergence in distribution. It is then easy to conclude the proof of the proposition. □

When $k=1$, this is the classical result for the scaled $M/M/\infty$ queue that, for the convergence in distribution

$$\lim_{N \to +\infty} \left(\frac{X_N(t)}{N} \right) = \left(\frac{\lambda}{\mu} + \left(\alpha - \frac{\lambda}{\mu} \right) e^{-\mu t} \right).$$

See Theorem 6.13 in Robert [67].

The case $k\geq 2$ is in fact more interesting, and more important for our study. With Definition (4.14), the above proposition gives the asymptotic behavior of the process $(\overline{X}^N(t/N^{1-1/k}))$, i.e. on a slower timescale than the timescale (t) of interest in our paper. It is quite clear that $(\overline{X}_N(t))$ should be close to the equilibrium of the function $(x(t))$, i.e. close to $\ell_\infty = \sqrt[k]{\lambda/\mu}$.

For such a process on a fast timescale, a convergence result of $(\overline{X}^N(t))$ to (ℓ_∞) is classically formulated in terms of the convergence in distribution of its *occupation measure*. See Section 2.3. Here, however, a stronger result of convergence is a key ingredient in the proofs of tightness for the convergence results of this paper.

PROPOSITION 4.6. *If $k\geq 2$ and $X_N(0)=O(\sqrt[k]{N})$, then for any $0<\eta<T$, and $\varepsilon>0$,*

$$\lim_{N\to+\infty} \mathbb{P}\left(\sup_{\eta\leq t\leq T} \left| \frac{X^N(t)}{\sqrt[k]{N}} - \ell_\infty \right| > \varepsilon \right) = 0$$

holds with $\ell_\infty \stackrel{\text{def.}}{=} \sqrt[k]{\lambda/\mu}$.

PROOF. The proof is carried out in two (similar) steps: with a stochastic upper bound of $\overline{X}^N(t)-\ell_\infty$, and then, with a stochastic lower bound of $\ell_\infty-\overline{X}^N(t)$.

First, we show that the process reaches the neighborhood of ℓ_∞ before time $\eta>0$ with high probability. Let $\ell_1>\ell_\infty$, define

$$S_N \stackrel{\text{def.}}{=} \inf\left\{ t\geq 0 : X_N(t)^{(k)} \leq (\ell_1)^k N \right\},$$

The integration of Relation (4.22) gives

$$\begin{aligned}\mathbb{E}\left(X^N(\eta\wedge S_N)\right) &= x_N + k\mathbb{E}\left(\int_0^{\eta\wedge S_N} (\lambda N - \mu(X^N(u))^{(k)})\,\mathrm{d}u \right) \\ &\leq C_0\sqrt[k]{N} + k\mu\left((\ell_\infty)^k - (\ell_1)^k\right) N\mathbb{E}(\eta\wedge\tau_N),\end{aligned}$$

for some constant C_0. Therefore we have for N large enough,

$$\mathbb{E}\left(\eta\wedge S_N\right) \leq \frac{C_0}{k\mu((\ell_1)^k-(\ell_\infty)^k)} N^{1/k-1},$$

and therefore that $(\mathbb{P}(S_N>\eta))$ converges to 0.

With the strong Markov property of $(X_N(t))$, we can therefore assume that

$$X_N(0) \leq y_N \stackrel{\text{def.}}{=} \ell_1\sqrt[k]{N} + k - 1.$$

Let $(Z(t))$ be a birth and death process on $\mathbb{N}$ starting at 0, with the transitions

$$x \to x + \begin{cases} +1 & \lambda, \\ -1 & \mu(\ell_1)^k \text{ if } x\geq 1. \end{cases}$$

The process $(Z(t))$ is the process of the number of jobs of an $M/M/1$ queue with input rate λ and service rate $\mu(\ell_1)^k$. See Chapter 5 of Robert [67]. Since $\mu(\ell_1)^k>\lambda$, this process is positive recurrent.

We now construct a coupling of $(X_N(t))$ and $(Z(t))$ such that the relation

$$X_N(t) \leq y_N + kZ(Nt), \quad \forall t\geq 0, \tag{4.24}$$

holds, where $(X_N(t))$ is the solution of the SDE (4.22) with initial point $x_N\leq y_N$ and $(Z(t))$ is the solution of the SDE

$$\mathrm{d}Z(t) = \mathcal{P}_{01}\left((0,\lambda N), \frac{\mathrm{d}t}{N}\right) - \mathbb{1}_{\{Z(t-)>0\}}\mathcal{P}_{10}\left((0,\mu(\ell_1)^k N), \frac{\mathrm{d}t}{N}\right)$$

with initial point at 0.

It is enough to prove Relation (4.24) by induction on the instants of jumps of the process $(X_N(t), Z(Nt))$ in the following way: if the inequality holds at time t_0, then it also holds at the instant of the next jump of the process $(X_N(t), Z(Nt))$ after time t_0.

Without loss of generality, we can assume that $t_0{=}0$ and $X_N(0){\le}y_N{+}kZ(0)$, t_1 is the first instant of jump of $(X_N(t), Z(Nt))$. Since both processes $(X^N(t))$ and $(kZ(Nt))$ have the same positive jump sizes at the same instants, we have only to consider jumps with negative sizes.

(a) If $X_N(0){\ge}y_N$, then $X_N(0)^{(k)}{\ge}(\ell_1)^k$. If at time t_1, there is a jump for $(Z(Nt))$ whose size is $-k$, it is due to the Poisson process $\mathcal{P}_{10}$. In view of the SDE for $(X_N(t))$, this implies that there is also a jump $-k$ for $(X_N(t))$ at time t_1. Relation (4.24) will then also hold at the first instant of jump of $(X_N(t), Z(Nt))$.

(b) If $X_N(0){<}y_N$, if there is a negative jump of $(Z_N(Nt))$ at time t_1, Relation (4.24) will obviously hold at that instant.

All the other possibilities preserve clearly the desired inequality.

Now, for ℓ_2 such that $\ell_2{>}\ell_1$,

$$\mathbb{P}\left(\sup_{0\le t\le T}\frac{X^N(t)}{\sqrt[k]{N}}\ge\ell_2\right)\le\mathbb{P}\left(\sup_{0\le t\le T}Z(Nt)\ge\frac{(\ell_2{-}\ell_1)\sqrt[k]{N}}{k}{-}1\right).$$

If, for $0{<}\varepsilon{<}\ell_2{-}\ell_1$,

$$\tau_N\stackrel{\text{def.}}{=}\inf\{t\ge0:Z(t)\ge\varepsilon\sqrt[k]{N}\},$$

with the last inequality, we have therefore, for N sufficiently large,

$$\mathbb{P}\left(\sup_{0\le t\le T}\frac{X^N(t)}{\sqrt[k]{N}}\ge\ell_2\right)\le\mathbb{P}(\tau_N\le NT). \tag{4.25}$$

Proposition 5.11 of Robert [67] on the hitting times of a positive recurrent $M/M/1$ queue gives that there exists $\rho{\in}(0,1)$ such that the sequence $(\rho^{\sqrt[k]{N}}\tau_N)$ converges in distribution to an exponentially distributed random variable. In particular

$$\limsup_{N\to+\infty}\mathbb{P}(\tau_N\le TN)=\limsup_{N\to+\infty}\mathbb{P}\left(\rho^{\sqrt[k]{N}}\tau_N\le N\rho^{\sqrt[k]{N}}T\right)=0.$$

Since ℓ_2 is arbitrarily close to ℓ_∞, Relation 4.25 gives the relation for the upper bound. The other case uses the same ingredients. The proposition is proved. □

With the same type of arguments, we can obtain the following corollary.

COROLLARY 4.7. *For $k{\ge}2$, if the initial condition of $(X_N(t))$ satisfies the relation*

$$\lim_{N\to+\infty}\frac{X_N(0)}{\sqrt[k]{N}}=\alpha>0,$$

and, for $M{>}m{>}0$ such that α^k, $\lambda/\mu{\in}(m,M)$. then

$$\lim_{N\to+\infty}\mathbb{P}\left(\frac{X^N(s)^{(k)}}{N}\in(m,M),\forall s{\in}[0,T]\right)=1.$$

4. Uniform Estimates

This section is devoted to the proof of the fact that for any $T{>}0$, with high probability, the scaled process $(\overline{X}_N(t))$ of Relation (4.14) lives in a bounded domain of the interior of $\mathbb{R}^n_+$ uniformly on the time interval $[0,T]$. Recall that since the components with index i such that $k_i{\ge}2$ are on "fast" timescales, see Section 2.6, uniform estimates on a time interval are more challenging to establish.

THEOREM 4.8. *If $(X_N(t))$ is the Markov process associated to the k-unary CRN of Definition 2.1 whose matrix R_κ is irreducible and with initial conditions satisfying Relation (4.15), then for any $T>0$, there exist two positive vectors (m_i) and (M_i) of $\mathbb{R}_+^n$ such that*

$$\lim_{N\to+\infty}\mathbb{P}\left(\mathcal{E}_N\right)=\lim_{N\to+\infty}\mathbb{P}\left(\frac{X_i^N(t)^{(k_i)}}{N}\in(m_i,M_i),\forall i\in\{1,\dots,n\},\forall t\leq T\right)=1,$$

with, for $J\subset I$,

$$\begin{cases}\mathcal{K}_J= & \left\{x=(x_i)\in(\mathbb{R}_+)^{I^*}\,:\,\sqrt[k_i]{m_i}<(x_i)^{k_i}<\sqrt[k_i]{M_i},\forall i\in J^*\right\}\\ \mathcal{E}_N\stackrel{\text{def.}}{=} & \left\{\overline{X}_N(t)\in\mathcal{K}_I,\forall t\leq T\right\}.\end{cases}\tag{4.26}$$

The important implication of this result is that, on the set $\mathcal{E}_N$, *every reaction* has a rate of the order of N. Note that because of the factorial term in the expression of the rate of the reactions, the event $\mathcal{E}_N$ is not the event

$$\left\{\frac{X_i^N(t)^{(k_i)}}{N}\in(m_i,M_i),\forall i\in\{1,\dots,n\},\forall t\leq T\right\},$$

however, when N goes to infinity, both events have the same probability.

The proof of the theorem is done by considering the stopping time $H_N\wedge T_N$, where

$$\begin{cases}H_N\stackrel{\text{def.}}{=}\inf\left\{t\geq 0:\min\limits_{i\in I^*}\dfrac{(X_i^N(t))^{(k_i)}}{m_iN}\leq 1\right\},\\ T_N\stackrel{\text{def.}}{=}\inf\left\{t\geq 0:\max\limits_{i\in I^*}\dfrac{(X_i^N(t))^{(k_i)}}{M_iN}\geq 1\right\},\end{cases}\tag{4.27}$$

and prove that for any $T>0$, the sequence $(\mathbb{P}(H_N\wedge T_N\leq T))$ converges to 0.

The proof is done in several steps. Results on convenient vectors (m_i) and (M_i) are established in Section 4.1. Proposition 4.11 of Section 4.2 proves the result when $I^*_{[1]}=\emptyset$, i.e. when $k_i\geq 2$ for all $i\in I^*$. Section 4.3 concludes with the general case. A coupling argument with a set of independent $M/M/\infty$ queues is used and then Corollary 4.7 of Section 3.

4.1. Some Linear Algebra. The notations and assumptions of Section 2.3 are used.

PROPOSITION 4.9. *If $\kappa\in\Omega(I)$, then there exists a unique solution $\ell_\kappa=(\ell_{\kappa,i})\in(\mathbb{R}_+^*)^{I^*}$, such that, for $i\in I^*$,*

$$\kappa_i^+(\ell_{\kappa,i})^{k_i}=\kappa_{0i}+\sum_{j\in I^*\setminus\{i\}}(\ell_{\kappa,j})^{k_j}\kappa_{ji},\tag{4.28}$$

furthermore,

$$((\ell_{\kappa,i})^{k_i})=M_\kappa^R\cdot\left(\frac{\kappa_{0i}}{\kappa_i^+}\right),\tag{4.29}$$

where M_κ^R is an $I^*\times I^*$ matrix whose coefficients are non-negative and depend only on κ_{ij}, $i\in I^*$, $j\in I$.

Recall that, from Relation (4.8), if $i\in I^*$,

$$\kappa_i^+=\kappa_{i0}+\sum_{j\neq i}\kappa_{ij}.$$

PROOF. The system (4.28) can be written as

$$z \cdot R_\kappa = 0,$$

with z_0=1 and z_i=$(\ell_{\kappa,i})^{k_i}$, for $i\in I^*$. This is simply the system of invariant measure equations for the Markov process associated to R_κ, introduced in Section 2.1. Since I is finite, the irreducibility property gives the existence and uniqueness of such a solution z.

Relation (4.29) is just a linear algebra representation of this solution, based on the fact that the spectral radius of the matrix R^*_κ=$(\kappa_{ji}/\kappa_i^+, i, j\in I^*)$ is strictly less than 1, which is a consequence of the irreducibility of R_κ,

$$\left(\kappa_i^+(\ell_{\kappa,i})^{k_i}\right) = \left(\sum_{m=0}^{+\infty}(R^*_\kappa)^m\right)\cdot(\kappa_{0i}).$$

□

The following proposition is a key result used in a coupling in the proof of Theorem 4.8.

PROPOSITION 4.10. *If $\kappa\in\Omega(I)$, then for any $(\alpha_i)\in(\mathbb{R}_+^*)^n$, there exist two vectors (m_i) and $(M_i)\in\mathbb{R}_+^n$ such that,*

$$0 < m_i < \alpha_i^{k_i} < M_i, \quad \forall i \in I^*, \tag{4.30}$$

and

$$M_i\kappa_i^+ > \kappa_{0i} + \sum_{j\in I^*\setminus\{i\}} M_j\kappa_{ji}, \tag{4.31}$$

$$m_i\kappa_i^+ < \kappa_{0i} + \sum_{j\in I^*,\ d(j)<d(i)} m_j\kappa_{ji}, \tag{4.32}$$

where d is the distance of Definition 4.2.

It should be noted that Relations (4.31) and (4.32) are not symmetrical, because of the restriction on the summation using the distance d. The result will used for the vector (α_i) associated to the initial conditions, see Relation (4.15).

PROOF. Let (z_i)=$((\ell_{\kappa,i})^{k_i})$ be the solution of the system of the type (4.28), for i=1,..., n,

$$z_i\kappa_i^+ = 1 + \sum_{j\in I^*\setminus\{i\}} z_j\kappa_{ji}.$$

Relations $\alpha_i^{k_i}$<M_i and (4.31) hold if we take M_i=ρz_j, with

$$\rho > \max\left(\frac{\alpha_i^{k_i}}{z_i}, \kappa_{0i} : i=1,\dots,n,\right).$$

The construction of (m_i) for the lower bounds is done by induction on the values of $d(i)$.

If $i\in I^*$ is such that $d(i)$=1, then necessarily κ_{0i}>0, then we can take m_i so that

$$0 < m_i < \min\left(\frac{\kappa_{0i}}{\kappa_i^+}, \alpha_i^{k_i}\right).$$

If $d(i)$=$p\geq$2, then there exists $j\in I^*$ such that $d(j)$=$p-1$ and κ_{ji}>0, therefore we can take m_i>0 such that

$$m_i < \min\left(\alpha_i^{k_i}, \frac{1}{\kappa_i^+}\left(\kappa_{0i} + \sum_{j\in I^*, d(j)<d(i)} m_j\kappa_{ji}\right)\right),$$

since the sum of the second term is strictly positive. The proposition is proved. □

4.2. CRN with Only Fast Processes. In this section it is assumed that $I^*_{[1]}$ is empty, i.e. $k_i \geq 2$ for all $i \in \{1, \dots, n\}$.

PROPOSITION 4.11. *If $I^*_{[1]}=\emptyset$, then there exist two vectors (m_i) and (M_i) with positive coordinates such that for any $T>0$, the sequence $(\mathbb{P}(\mathcal{E}_N))$ is converging to 1, where $\mathcal{E}_N$ is the event defined by Relation (4.26).*

PROOF. Let H_N and T_N be the stopping times defined by Relation (4.27). We start with the stopping time T_N. We take (M_i) of Proposition 4.10 satisfying Relations (4.30) and (4.31). Let $(Y_N(t)) = (Y_i^N(t))$ be the solution of the SDE

$$\mathrm{d}Y_i^N(t) = k_i \mathcal{P}_{0i}\left((0, \kappa_{0i}N), \mathrm{d}t\right) + \sum_{j \in I^* \setminus \{i\}} k_i \mathcal{P}_{ji}\left((0, \kappa_{ji} M_j N), \mathrm{d}t\right) - \sum_{j \in I \setminus \{i\}} k_i \mathcal{P}_{ij}\left(\left(0, \kappa_{ij} (Y_i^N(t-))^{(k_i)}\right), \mathrm{d}t\right),$$

with $Y_N(0)=X_N(0)$. Note that we have necessarily that $Y_i^N(t)-X_i^N(t) \in k_i\mathbb{Z}$, for all $i \in I^*$ and $t \geq 0$.

We prove that, for all $t<T_N$, the relations $X_i^N(t) \leq Y_i^N(t)$ hold for any $i \in I^*$. This is done by induction on the sequence of the instants of jumps of the process $((X_i^N(t), Y_i^N(t)), i \in I^*)$ in the time interval $[0, T_N]$. As in the proof of Proposition 4.6, we assume that $X_i^N(0) \leq Y_i^N(0)$ and $X_i^N(0)^{(k_i)} \leq M_i N$ hold for all $i \in I^*$, and denote by t_1 the instant of the first jump of the process $(X_i^N(t), Y_i^N(t), i \in I^*)$. We show that the above inequalities also hold at time t_1.

For all $i \in I^*$, we have $(X_i^N(0))^{(k_i)} \leq M_i N$, hence, for, $j \in I^*$ and $t \geq 0$,

$$\mathcal{P}_{ji}\left(\left(0, \kappa_{ji} (X_i^N(t))^{(k_i)}\right) \times [0, t]\right) \leq \mathcal{P}_{ji}\left((0, \kappa_{ji} M_j N) \times [0, t]\right),$$

and if $(X_i^N(t))$ has a jump up at t_1 due to $\mathcal{P}_{ji}$, so does $(Y_i^N(t))$. Consequently, the inequality is clearly preserved if the size of the first jump is positive.

If t_1 is an instant of a jump with negative size for $(Y_i^N(t))$, if $X_i^N(0) < Y_i^N(0)$, then necessarily $Y_i^N(0) - X_i^N(0) \geq k_i$, the relation $X_i^N(t_1) \leq Y_i^N(t_1)$ is therefore satisfied. All the other possibilities for t_1 clearly preserve the desired relations. Our assertion has been established.

For $i \in I^*$, the process $(Y_i^N(t))$ has the same distribution as the process of a generalized $M/M/\infty$ queue, introduced in Section 3, with arrival rate $\lambda_i N$ and departure rate μ_i given by

$$\lambda_i = \kappa_{0i} + \sum_{j \in I^* \setminus \{i\}} M_j \kappa_{ji}, \quad \mu_j = \kappa_i^+.$$

We have $M_i > \lambda_i/\mu_i$ for all $i \in I^*$ because of Relation (4.31). Since all k_i's are greater than 2, Corollary 4.7 applied to these n generalized $M/M/\infty$ queues shows that the relation

$$\lim_{N \to +\infty} \mathbb{P}(T_N \leq T) = 0$$

holds. We now take care of the stopping time H_N. A vector (m_i) satisfying Relations (4.30) and (4.32) of Proposition 4.10 is fixed. Let $(Z_N(t)) = (Z_i^N(t))$ be the solution of the SDE

$$\mathrm{d}Z_i^N(t) = k_i \mathcal{P}_{0i}\left((0, \kappa_{0i}N), \mathrm{d}t\right) + \sum_{\substack{j \in I^* \\ d(j) < d(i)}} k_i \mathcal{P}_{ji}\left((0, \kappa_{ji} m_j N), \mathrm{d}t\right) - \sum_{j \in I \setminus \{i\}} k_i \mathcal{P}_{ij}\left(\left(0, \kappa_{ij} (Y_i^N(t-))^{(k_i)}\right), \mathrm{d}t\right),$$

with $Z_N(0)=X_N(0)$. It is easily seen by induction on the sequence of the instants of jumps of the process $(X_i^N(t), Z_i^N(t))$ that the relation $X_i^N(t)\geq Z_i^N(t)$ holds for all $t<H_N$ and $i\in I^*$.

For $i\in I^*$, the process $(Z_i^N(t))$ has the same distribution as the process of a generalized $M/M/\infty$ queue with arrival rate $\lambda_i N$ and departure rate μ_i given by

$$\lambda_i \stackrel{\text{def.}}{=} \kappa_{0i} + \sum_{\substack{j\in I^* \\ d(j)<d(i)}} m_j \kappa_{ji}, \quad \mu_j \stackrel{\text{def.}}{=} \kappa_i^+.$$

Since the vector (m_i) has been chosen so that $m_i<\lambda_i/\mu_i$ holds for all $i\in I^*$, we can conclude in the same way as before using Corollary 4.7. The proposition is proved. □

4.3. Proof of Theorem 4.8. We first take care of the indices in the set $I^*_{[1]}$. We define

$$\begin{cases} m_1^1 = \dfrac{1}{2}\alpha_{\min}\exp(-\kappa^+_{\max}T), \\ M_1^1 = 2k_{\max}\left(\kappa_0^+ T + \displaystyle\sum_{j\in I^*_{[1]}} \alpha_j\right), \end{cases} \tag{4.33}$$

with $x_{\max/\min} = \max/\min(x_i, 1\leq i\leq n)$ for $x\in\mathbb{R}^n_+$.

We show here that for all $i\in I^*_{[1]}$, we can choose $m_i=m_1^1$ and $M_i=M_1^1$. For all $i\in I^*$, it is easily seen that the following upper bound, for $t\geq 0$,

$$\sup_{t\leq T}\sum_{i\in I^*_{[1]}} k_i X_i^N(t) \leq k_{\max}\left(\sum_{i\in I^*_{[1]}} x_i^N + \sum_{i\in I^*_{[2+]}} x_i^N + \sum_{i\in I^*_{[1]}} \mathcal{P}_{0i}([0,\kappa_{0i}N]\times[0,T])\right). \tag{4.34}$$

holds. The right-hand side of the last relation divided by N converges almost surely to

$$k_{\max}\left(\kappa_0^+ T + \sum_{i\in I^*_{[1]}} \alpha_i\right),$$

hence

$$\lim_{N\to+\infty} \mathbb{P}\left(\sup_{t\leq T}\max_{i\in I^*_{[1]}} \frac{X_i^N(t)}{N} \geq M_1^1\right) = 0. \tag{4.35}$$

Since the lifetime of a molecule of type $i\in I^*_{[1]}$ is exponentially distributed with parameter κ_i^+, the number of species i at time T is stochastically greater than

$$\sum_{k=1}^{x_i^N} \mathbb{1}_{\{E_k^{i+}\geq T\}},$$

where (E_k^{i+}) is a sequence of i.i.d. exponential random variables with parameter κ_i^+. This last quantity divided by N converges almost surely to $\alpha_i\exp(-\kappa_i^+ T)$. We therefore obtain the relation

$$\lim_{N\to+\infty} \mathbb{P}\left(\inf_{t\leq T}\min_{i\in I^*_{[1]}} \frac{X_i^N(t)}{N} \leq m_1^1\right) = 0. \tag{4.36}$$

From Relations (4.35) and (4.36), for node $i\in I^*_{[2+]}$, the input rate from node $j\in I^+_{[1]}$ on the time interval $[0,T]$ is, with high probability, upper bounded by $\kappa_{ji}M_1^1$ and lower bounded by $\kappa_{ji}m_1^1$.

Define $\overline{\kappa}=(\overline{\kappa}_{ij}, i,j\in I_{[2+]})$ and $\underline{\kappa}=(\underline{\kappa}_{ij}, i,j\in I_{[2+]})$, by, for $i\in I_{[2+]}$,

$$\begin{cases} \overline{\kappa}_{ij}=\underline{\kappa}_{ij}=\kappa_{ij}, \quad j\in I_{[2+]};\\ \overline{\kappa}_{i0}=\underline{\kappa}_{i0}=\kappa_{i0}+\sum_{j\in I^*_{[1]}}\kappa_{ij};\\ \overline{\kappa}_{0i}=\kappa_{0i}+\sum_{j\in I^*_{[1]}}\kappa_{ji}M^1_1;\\ \underline{\kappa}_{0i}=\kappa_{0i}+\sum_{j\in I^*_{[1]}}\kappa_{ji}m^1_1. \end{cases}$$

Using a coupling argument, one can define the Markov processes $(Y^2_N(t))$, respectively $(Z^2_N(t))$, associated to the k-Unary CRN with species $I^*_{[2,+]}$, with complexes $(k_iS_i, i\in I^*_{[2+]})$ and constant of reactions $\overline{\kappa}$, respectively $\underline{\kappa}$, both starting at $X^N_{[2,+]}(0)$ and that verify for all $t\le T_N\wedge H_N$,

$$Z^{N,2}_i(t)\le X^N_i(t)\le Y^{N,2}_i(t), \quad \forall i\in I^*_{[2+]}.$$

Since $\overline{\kappa}\in\Omega(I_{[2+]})$, Proposition 4.11 applied to the process $(Y^2_N(t))$ shows that there exists a vector $(M_i, \in I_{[2+]})$, such that

$$\lim_{N\to+\infty}\mathbb{P}\left(\overline{X}_N(t)\in\prod_{i=1}^N\left(0,\ \sqrt[k_i]{M_i}\right), \forall t\le T\right)=1.$$

Similarly, by considering $\underline{\kappa}$, there exists a vector $(m_i, \in I_{[2+]})$ with positive components such that

$$\lim_{N\to+\infty}\mathbb{P}\left(\overline{X}_N(t)\in\prod_{i=1}^N\left(\sqrt[k_i]{m_i},\ \sqrt[k_i]{M_i}\right), \forall t\le T\right)=1.$$

The theorem is proved.

5. CRN with only fast processes

When $I^*_{[1]}$ is empty, i.e. $k_i\ge 2$ for all $i\in\{1,\dots,n\}$, the time evolutions of all species are fast processes, see Section 2.6. Theorem 4.4 is only about the convergence in distribution of the sequence of occupation measures (Λ_N) on $\mathbb{R}_+\times\mathbb{R}^n_+$ defined by Relation (4.16). The absence of chemical species i such that $k_i=1$ gives a kind of instantaneous equilibrium property in the sense that the limit in distribution of (Λ_N) is homogeneous with respect to the first coordinate, the time coordinate. The main result of this section is Theorem 4.17 which is simply Theorem 4.4 stated in this context. The motivation of such a separate proof is that it is focused, in our view, on the key argument of the general proof. The identification of possible limits of (Λ_N) is done by induction via the use of an entropy function. The proof of the general case follows also such line but in a "non-homogeneous", technically more complicated, context.

5.1. Tightness of (Λ_N). We first establish the tightness of (Λ_N) for the convergence in distribution in the general case.

PROPOSITION 4.12. *If the subset $I^*_{[1]}$ is empty and if the initial conditions satisfy Relation (4.15), then the sequence of measure valued processes (Λ_N) on $[0,T]\times(\mathbb{R}^*_+)^{I^*}$ is tight for the convergence in distribution. Any limiting point Λ_∞ can be expressed as,*

$$\langle\Lambda_\infty, f\rangle=\int_{[0,T]\times\mathcal{K}_I}f(s,x)\pi_s(\mathrm{d}x)\,\mathrm{d}s, \tag{4.37}$$

for any function $f\in\mathcal{C}_c([0,T]\times(\mathbb{R}_+)^{I^*})$, where (π_s) is an optional process with values in $\mathcal{P}(\mathcal{K}_I)$, the set of probability measures on the compact subset $\mathcal{K}_I$ defined by Relation (4.26).

See Dawson [21] for a presentation of the convergence in distribution of measure-valued processes. The optional property of (π_s) is used only to have convenient measurability properties so that time-integrals with respect to $(\pi_s, s{>}0)$ are indeed random variables. See Section VI.4 of Rogers and Williams [69].

PROOF. We take the vectors (m_i) and (M_i) of Theorem 4.8, and $\mathcal{K}_I$ the compact set of $\mathbb{R}_+^n$ and $\mathcal{E}_N$ the event defined in Relation (4.26). Since $\Lambda_N([0,T]{\times}\mathcal{K}_I){\geq}T\mathbb{1}_{\mathcal{E}_N}$, with Theorem 4.8, we obtain the relation

$$\lim_{N\to+\infty} \mathbb{E}\left(\Lambda_N\left([0,T]{\times}\mathcal{K}_I\right)\right) = T.$$

Lemma 1.3 of Kurtz [55] gives that the sequence of random measures (Λ_N) is tight for the convergence in distribution, and Lemma 1.4 of the same reference gives the representation (4.37). The proposition is proved. □

In the following we assume that Λ_∞ is a limit of a subsequence (Λ_{N_r}) with the representation (4.37).

LEMMA 4.13. *If f is a continuous function on $\mathbb{R}_+^{I^*}$, then the relation*

$$\lim_{r\to+\infty}\left(\int_0^t f\left(\overline{X}_{N_r}(s)\right)\mathrm{d}s\right) = \left(\int_0^t\int_{\mathbb{R}_+^{I^*}} f(x)\pi_s(\mathrm{d}x)\,\mathrm{d}s,\right)$$

holds for the convergence in distribution of processes.

PROOF. This is a straightforward use of the criterion of modulus of continuity, see Theorem 7.3 of Billingsley [14], and of Theorem 4.8. For $s{\leq}t$, on the event $\mathcal{E}_N$, we have

$$\int_s^t f\left(\overline{X}_{N_r}(s)\right)\mathrm{d}s \leq 2(t{-}s)\sup_{x\in\mathcal{K}_I}|f(x)|,$$

with the notations of Relation (4.26). We conclude with the identification of the finite marginals. □

As we have seen in Section 2.6, for $i{\in}I^*$, the value of k_i gives in fact the natural timescale of the process $(\overline{X}_i^N(t))$. On the event $\mathcal{E}_N$, see Relation (4.26), every reaction has a rate of order N, in particular, the rate at which the process $(X_i^N(t))$ jumps of $\pm k_i$ is of order N. With the scaling in space of the process, $(\overline{X}_i^N(t))$ is significantly changed when there are N^{1/k_i} reactions changing $(X_i^N(t))$, and therefore after a duration of time of the order of N^{1/k_i-1}. If for two species i and j, $k_i{>}k_j$, then the process $(\overline{X}_i^N(t))$ changes more rapidly than the process $(\overline{X}_j^N(t))$.

From now on in this section it is assumed that $I^*_{[1]}$ is empty.

5.2. A Limiting Equation. For a function $f\in\mathcal{C}_c^2\left((\mathbb{R}_+)^{I^*}\right)$, the SDE (4.10) gives directly, for $t{\in}[0,T]$,

$$\begin{aligned}(4.38)\quad f\left(\overline{X}_N(t)\right) = f\left(\overline{X}_N(0)\right) + M_{f,N}(t) + \int_0^t \sum_{i\in I^*}\kappa_{0i}N\nabla_{\frac{k_i}{N^{1/k_i}}e_i}(f)(\overline{X}_N(s))\,\mathrm{d}s\\ + \int_0^t \sum_{\substack{i,j\in I,\\ i\neq 0}}\kappa_{ij}(X_i^N(s))^{(k_i)}\nabla_{-\frac{k_i}{N^{1/k_i}}e_i+\frac{k_j}{N^{1/k_j}}e_j}(f)(\overline{X}_N(s))\,\mathrm{d}s,\end{aligned}$$

with the notations

— for x, $a{\in}\mathbb{R}^{I^*}$, $\nabla_a(f)(x){=}f(x{+}a){-}f(x)$;
— for $i{\in}I^*$, e_i is the i-th unit vector of $\mathbb{R}^{I^*}$, and the convention $e_0{=}0$,

and $(M_{f,N}(t))$ is local martingale whose previsible increasing process is given by, for $t\leq T$,

$$\langle M_{f,N}\rangle(t) = \int_0^t \sum_{i\in I^*} \kappa_{0i} N \left(\nabla_{\frac{k_i}{N^{1/k_i}}e_i}(f)(\overline{X}_N(s))\right)^2 \mathrm{d}s + \int_0^t \sum_{\substack{i,j\in I,\\ i\neq 0}} \kappa_{ij}(X_i^N(t))^{(k_i)} \left(\nabla_{-\frac{k_i}{N^{1/k_i}}e_i+\frac{k_j}{N^{1/k_j}}e_j}(f)(\overline{X}_N(s))\right)^2 \mathrm{d}s \tag{4.39}$$

PROPOSITION 4.14. *If the subset $I^*_{[1]}$ is empty and (Λ_∞) is a limiting point of (Λ_N) with the representation (4.37), then, for any $p\geq 2$ and $f\in\mathcal{C}^2(\mathcal{K}_{I_{[2,p]}})$, almost surely, the relation*

$$\int_0^t \int_{\mathcal{K}_I} \sum_{i\in I^*_{[p]}} \left(\kappa_{0i} + \sum_{j\in I^*\setminus\{i\}} \kappa_{ji} x_j^{k_j} - \kappa_i^+ x_i^p\right) \frac{\partial f}{\partial x_i}(x_{[2,p]})\pi_s(\mathrm{d}x)\,\mathrm{d}s = 0, \tag{4.40}$$

holds for all $t\in[0,T]$.

Recall the conventions $x_{[2,p]}=(x_i, i\in I^*_{[2,p]})$ for $x\in(\mathbb{R}_+)^{I^*}$, see Section 2.3.

PROOF. It is assumed that $I^*_{[p]}\neq\emptyset$. Let $f\in\mathcal{C}^2((\mathbb{R}_+)^{I^*_{[2,p]}})$. To simplify expressions in this proof, we will make the slight abuse of notation, $f(x)=f(x_{[2,p]})$ for $x\in(\mathbb{R}_+)^{I^*}$.

Since our goal is of characterizing the process (π_t), by Theorem 4.8, without loss of generality, we can assume that the support of the function f is included in $\mathcal{K}_I$ defined in Relation (4.26). Similarly, from now on, all relations are considered on the event $\mathcal{E}_N$ whose probability is arbitrarily close to 1 as N gets large. In particular the process $(\overline{X}_N(t), t\in[0,T])$ has values in $\mathcal{K}_I$.

For $t\leq T$, Relation (4.38) can be rewritten as,

$$\begin{aligned}
&\frac{f\left(\overline{X}_N(t)\right)}{N^{1-1/p}} - \frac{f\left(\overline{X}_N(0)\right)}{N^{1-1/p}} - \frac{M_{f,N}(t)}{N^{1-1/p}} \\
&= \int_0^t \sum_{i\in I^*_{[2,p]}} \left(\kappa_{0i} + \sum_{j\notin I_{[2,p]}} \kappa_{ji}\frac{(X_j^N(t))^{(k_j)}}{N}\right) N^{1/p}\nabla_{\frac{k_i}{N^{1/k_i}}e_i}(f)(\overline{X}_N(s))\,\mathrm{d}s \\
&+ \int_0^t \sum_{i\in I^*_{[2,p]}} \left(\kappa_{i0} + \sum_{j\notin I_{[2,p]}} \kappa_{ij}\right) \frac{(X_i^N(t))^{(k_i)}}{N} N^{1/p}\nabla_{-\frac{k_i}{N^{1/k_i}}e_i}(f)(\overline{X}_N(s))\,\mathrm{d}s \\
&+ \int_0^t \sum_{i\in I^*_{[2,p]}} \sum_{j\in I^*_{[2,p]}\setminus\{i\}} \kappa_{ij}\frac{(X_i^N(t))^{(k_i)}}{N} N^{1/p}\nabla_{-\frac{k_i}{N^{1/k_i}}e_i+\frac{k_j}{N^{1/k_j}}e_j}(f)(\overline{X}_N(s))\,\mathrm{d}s.
\end{aligned} \tag{4.41}$$

For a, $b\geq 0$, there exist constants C_0 and C_1 such that

$$\max_{i\in I^*}\sup_{x\in\mathcal{K}_I}\left|x^{k_i} - \frac{(\sqrt[k_i]{N}x)^{(k_i)}}{N}\right| \leq \frac{C_0}{N^{1/k_i}}, \tag{4.42}$$

and, for any i, $j\in I^*$,

$$\sup_{x\in\mathcal{K}_I}\left|\nabla_{-\frac{a}{N^{1/k_i}}e_i+\frac{b}{N^{1/k_j}}e_j}(f)(x) + \frac{a}{N^{1/k_i}}\frac{\partial f}{\partial x_i}(x) - \frac{b}{N^{1/k_j}}\frac{\partial f}{\partial x_j}(x)\right| \leq C_1\left(\frac{a}{N^{1/k_i}}+\frac{b}{N^{1/k_j}}\right).$$

We get that, for $i\in I_{[2,p]}$, the processes

$$\left(N^{1/p}\nabla_{\pm\frac{k_j}{N^{1/k_i}}e_i}(f)(\overline{X}_N(t)), t\le T\right)$$

vanish if $k_i\neq p$. With the definition (4.27), Relation (4.39) and Doob's Inequality give that the martingale $(M_{f,N}(t\wedge T_N)/N^{1-1/p})$ converges in distribution to 0 and so $(M_{f,N}(t)/N^{1-1/p})$ by Theorem 4.8.

Relation (4.41) becomes

$$\int_0^t \sum_{i\in I^*_{[p]}}\left(\kappa_{0i}+\sum_{j\in I^*\setminus\{i\}}\kappa_{ji}\left(\overline{X}^N_j(t)\right)^{k_j}\right)p\frac{\partial f}{\partial x_i}(\overline{X}_N(s))\,\mathrm{d}s \\ -\int_0^t\sum_{i\in I^*_{[p]}}\kappa^+_i\left(\overline{X}^N_i(t)\right)^{k_j}p\frac{\partial f}{\partial x_i}(\overline{X}_N(s))\,\mathrm{d}s = U_N(t),$$

where $(U_N(t))$ is a process converging in distribution to 0. This relation can be written in terms of occupation measure Λ_N, it is easy to conclude the proof of the proposition with the help of Lemma 4.13. □

5.3. A Convex Function on $\mathcal{K}_I$.

Definition 4.15. *If $\kappa\in\Omega(I)$, the function F_κ is defined by, for $z=(z_i)\in\mathcal{K}_I$,*

$$F_\kappa(z)\overset{\text{def.}}{=}\sum_{i\in I^*}\left(\kappa^+_i z_i-\kappa_{0i}-\sum_{j\in I^*\setminus\{i\}}\kappa_{ji}z_j\right)\ln\left(\frac{z_i}{(\ell_{\kappa,i})^{k_i}}\right), \tag{4.43}$$

where $\mathcal{K}_I$ is defined by Relation (4.26) *and $\ell_\kappa=(\ell_{\kappa,i})\in\mathbb{R}_+^{I^*}$ is the unique solution of the system* (4.28) *of Proposition 4.9.*

Proposition 4.16. *The function F_κ is non-negative, strictly convex on $\mathcal{K}_I$, with a unique minimum 0 at $z=((\ell_{\kappa,i})^k)$, furthermore the mapping $(\kappa,z)\mapsto F_\kappa(z)$ is continuous on $\Omega(I)\times\mathcal{K}_I$.*

Proof. The existence and uniqueness of ℓ_κ, solution of a non-singular linear system, has been seen in Proposition 4.9. The continuity of $\kappa\mapsto\ell_\kappa$ on $\Omega(I)$ gives the continuity of $(\kappa,z)\mapsto F_\kappa(z)$.

We now calculate the Hessian matrix of F_κ. For $i\in I^*$, we have, for $z\in\mathcal{K}_I$,

$$\frac{\partial F_\kappa}{\partial z_i}(z)=\kappa^+_i\ln\left(\frac{z_i}{(\ell_{\kappa,i})^{k_i}}\right)+\frac{1}{z_i}\left(\kappa^+_i z_i-\kappa_{0i}-\sum_{m\in I^*\setminus\{i\}}\kappa_{mi}z_m\right) \\ -\sum_{m\in I^*\setminus\{i\}}\kappa_{im}\ln\left(\frac{z_m}{(\ell_{\kappa,m})^{k_m}}\right).$$

Relation (4.28) gives that this quantity is indeed null at $z=((\ell_{\kappa,m})^{k_m})$. For $j\in I^*$, $j\neq i$, we have the relation

$$\frac{\partial^2 F_\kappa}{(\partial z_i)^2}(z)=\frac{1}{z_i^2}\left(\kappa^+_i z_i+\kappa_{0i}+\sum_{m\in I^*\setminus\{i\}}z_m\kappa_{mi}\right),\quad \frac{\partial^2 F_\kappa}{\partial z_i\partial z_j}(z)=-\frac{\kappa_{ij}z_i+\kappa_{ji}z_j}{z_iz_j}.$$

Let $\mathcal{H}_\kappa(z)$ be the Hessian matrix of F_κ at $z{\in}\mathcal{K}_I$. For $u{=}(u_i){\in}\mathbb{R}^{I^*}$, with the notation $\gamma_{ij}{=}\kappa_{ij}z_i{+}\kappa_{ji}z_j$ for $i{\neq}j$, the associated quadratic form at u is given by

$$u^t F_\kappa(z)u = -\sum_{i\in I^*}\sum_{j\in I^*\setminus\{i\}}\gamma_{ij}\frac{u_iu_j}{z_iz_j}+\sum_{i\in I^*}(\kappa_{i0}z_i{+}\kappa_{0i})\frac{u_i^2}{z_i^2}+\sum_{i\in I^*}\sum_{j\in I^*\setminus\{i\}}\gamma_{ij}\frac{u_i^2}{z_i^2}$$
$$=\sum_{i\in I^*}(\kappa_{i0}z_i{+}\kappa_{0i})\frac{u_i^2}{z_i^2}+\frac{1}{2}\sum_{i\in I^*}\sum_{j\in I^*\setminus\{i\}}\gamma_{ij}\left(\frac{u_i}{z_i}{-}\frac{u_j}{z_j}\right)^2.$$

This last expression is positive for any non-zero element $u{=}(u_i){\in}\mathbb{R}^{I^*}$. The function F_κ is strictly convex. This concludes the proof of the proposition. □

5.4. Identification of the Limit. We can now state the main convergence result of this section.

THEOREM 4.17. *If $\kappa{\in}\Omega(I)$ and the subset $I^*_{[1]}$ is empty, if Relation* (4.2) *holds for the initial conditions, then the sequence (Λ_N) is converging in distribution to Λ_∞, such that, almost surely, for any function $f{\in}\mathcal{C}_c(\mathbb{R}_+{\times}(\mathbb{R}_+^*)^{I^*})$, the relation*

$$\int f(s,x)\Lambda_\infty(\mathrm{d}s,\mathrm{d}x)=\int_0^{+\infty}f(s,\ell_\kappa)\,\mathrm{d}s, \tag{4.44}$$

holds, where $\ell_\kappa{=}(\ell_{\kappa,i})$ is the unique solution of the system (4.28) *of Proposition 4.9.*

The proof is carried out by induction on the "speed" of the different processes. We start by the identification of the fastest species, with the largest k_i, and identify step by step each set $I^*_{[p]}$. One of the difficulties is that we have only the functional equation, Relation (4.40), to identify all the species in the set $I^*_{[p]}$ for each $p\geq 2$. A convex function, related to a relative entropy functional, will be used to identify them simultaneously.

PROOF. Let $m_0{\geq}1$ and $(p_a){\in}\mathbb{N}^m$ such that $2{\leq}p_{m_0}{<}\cdots{<}p_2{<}p_1$ and

$$\{k_i, i{\in}I^*\}=\{p_a, a{=}1,\ldots,m_0\},$$

in particular, we have

$$I^*=\bigcup_{a=1}^{m_0}I^*_{[p_a]}\text{ and }I{=}I_{[2,p_1]}.$$

We will proceed by induction on m_0 to prove that a random measure Λ_∞ that verifies Relation (4.40) is expressed by Relation (4.44).

We first consider the species of the set $I^*_{[p_1]}$ associated to the fastest processes of $(X_N(t))$. With the notations of Relation (4.26), Relation (4.40) gives, for $T{>}0$ and p_1, the identity

$$\int_0^T\int_{\mathcal{K}_I}\sum_{i\in I^*_{[p_1]}}K_i\left[x_{[2,p_2]}\right](x_{[p_1]^{p_1}})\frac{\partial f}{\partial x_i}(x)\pi_s(\mathrm{d}x)\,\mathrm{d}s=0 \tag{4.45}$$

holds almost surely for $f{\in}\mathcal{C}^2(\mathcal{K}_I)$, with, for $y{\in}\mathcal{K}_{I_{[2,p_2]}}$, $z{\in}\mathcal{K}_{I_{[p_1]}}$ and $i{\in}I^*_{[p_1]}$,

$$K_i[y](z)\overset{\text{def.}}{=}\kappa_{0i}+\sum_{j\in I^*_{[2,p_2]}}y_j^{k_j}\kappa_{ji}+\sum_{j\in I^*_{[p_1]}\setminus\{i\}}z_j\kappa_{ji}-\kappa_i^+z_i.$$

and the notation $z^{p_1}{=}(z_i{}^{p_1})$.

For $y{\in}\mathcal{K}_{I_{[2,p_2]}}$, we introduce an $I_{[p_1]}{\times}I_{[p_1]}$ matrix $\overline{\kappa}^1(y)$ as follows: For $i,\,j{\in}I^*_{[p_1]}$, $j{\neq}i$, $\overline{\kappa}^1_{ij}(y){=}\kappa_{ij}$ and

$$\overline{\kappa}^1_{0i}(y)=\kappa_{0i}+\sum_{j\in I^*_{[2,p_2]}}y_j^{k_j}\kappa_{ji},\quad\overline{\kappa}^1_{i0}(y)=\kappa_{i0}+\sum_{j\in I^*_{[2,p_2]}}\kappa_{ij}.$$

Remark that, for $i \in I_{[p_1]}$,

$$\overline{\kappa}_i^{1,+}(y) = \sum_{j \in I_{[p_1]} \setminus \{i\}} \overline{\kappa}_{ij}^1(y) = \kappa_1^+.$$

It is easily seen that $\overline{\kappa}^1 \in \Omega(I_{[p_1]})$ and

$$K_i[y](z) = \overline{\kappa}_{0i}^1(y) + \sum_{j \in I^*_{[p_1]} \setminus \{i\}} z_j \overline{\kappa}_{ji}^1(y) - \overline{\kappa}_i^{1,+}(y) z_i.$$

Note that if $I^*_{[2,p_2]}$ is empty, then $\overline{\kappa}^1$ is then constant, there is no dependence on y of course, and Theorem 4.17 is proved for $m_0 = 1$.

Now if $I^*_{[2,p_2]}$ is not empty, for $y \in \mathcal{K}_{I_{[2,p_2]}}$, the equation

$$K_1[y](z^p) = 0$$

is the system (4.28) of Proposition 4.9 for the set of indices $I_{[p_1]}$ and the matrix $\overline{\kappa}^1(y)$. It has a unique solution $z = \widetilde{L}_1[y] = (L_i^1(y), i \in I^*_{[p_1]})$. We now define an *entropy function* H_1 given by, for $y \in \mathcal{K}_{I_{[2,p_2]}}$ and $z \in \mathcal{K}_{I_{[p_1]}}$,

$$H_1[y](z) = \sum_{i \in I^*_{[p_1]}} z_i \ln \left(\frac{z_i}{L_i^1(y)^{p_1}} \right) - z_i. \tag{4.46}$$

Note that $H_1[y]$ is a $\mathcal{C}^2$-function on $\mathcal{K}_I$. It is easily checked that Relation (4.45) for the function $f : x \mapsto H_1[x_{[2,p_2]}](x_{[p_1]})$ can be rewritten as

$$\int_0^T \int_{\mathcal{K}_I} F_1[x_{[2,p_2]}](x_{[p_1]}{}^{p_1}) \pi_s(\mathrm{d}x)\, \mathrm{d}s = 0, \tag{4.47}$$

where, for $z \in \mathcal{K}_{I_{[p_1]}}$,

$$F_1[y](z) \stackrel{\text{def.}}{=} \sum_{i \in I^*_{[p_1]}} \left(z_i \overline{\kappa}_i^{1,+}(y) - \overline{\kappa}_{0i}^1(y) - \sum_{j \in I^*_{[p_1]} \setminus \{i\}} z_j \overline{\kappa}_{ji}(y) \right) \ln \left(\frac{z_i}{L_i^1(y)^{p_1}} \right).$$

Note that, for $y \in \mathcal{K}_{I_{[2,p_2]}}$, $F_1[y]$ is the function $F_{\overline{\kappa}^1(y)}$ of Relation (4.43) for the set of indices $I_{[p_1]}$. Relation (4.47) gives therefore that, almost surely,

$$\int_0^T \int_{y \in \mathcal{K}_{I_{[2,p_2]}}} \left(\int_{z \in \mathcal{K}_{I_{[p_1]}}} F_1[y](z^{p_1}) \pi_s^{[p_1]}(\mathrm{d}z|y) \right) \mathrm{d}s \otimes \pi_s^{[2,p_2]}(\mathrm{d}y) = 0,$$

with the notations of Section 2.3 and, for $s \geq 0$, $\pi_s^{[p_1]}(\mathrm{d}z|y)$ is the conditional distribution on $\mathcal{K}_{I_{[p_1]}}$ of $\pi_s \in \mathcal{P}(\mathbb{R}_+^{I^*})$ with respect to $y \in \mathcal{K}_{I_{[2,p_2]}}$. Consequently, since $F_1[y]$ is non-negative, up to a negligible set of $[0,T] \times \mathcal{K}_{I_{[2,p_2]}}$ for the measure $\mathrm{d}s \otimes \pi_s^{[2,p_2]}(\mathrm{d}y)$, we have the relation

$$\int_{\mathcal{K}_{I_{[p_1]}}} F_1[y](z^{p_1}) \pi_s^{[p_1]}(\mathrm{d}z|y) = 0.$$

Proposition 4.16 gives that $\widetilde{L}_1(y)$ is the only root of the function $x \mapsto F_1[y](x^{p_1})$ on $\mathcal{K}_{I_{[p_1]}}$, hence the probability distribution $\pi_s^{[p_1]}(\mathrm{d}x|y)$ is the Dirac measure at $\widetilde{L}_1(y)$.

If h, f_1 and f_2, are continuous functions on, respectively, $[0,T]$, $\mathcal{K}_{I_{[2,p_2]}}$ and $\mathcal{K}_{I_{[p_1]}}$ then, almost surely,

$$\begin{aligned}\int_0^T \int_{\mathcal{K}_I} & h(s) f_1(x_{[2,p_2]}) f_2(x_{[p_1]}) \pi_s(\mathrm{d}x)\,\mathrm{d}s \\ &= \int_0^T \int_{y\in\mathcal{K}_{I_{[2,p_2]}}} f_1(y) \int_{z\in\mathcal{K}_{I_{[p_1]}}} h(s) f_2(z) \pi_s^{[p_1]}(\mathrm{d}z|y) \pi_s^{[2,p_2]}(\mathrm{d}y)\,\mathrm{d}s \\ &= \int_0^T \int_{y\in\mathcal{K}_{I_{[2,p_2]}}} h(s) f_1(y) f_2(\widetilde{L}_1(y)) \pi_s^{[2,p_2]}(\mathrm{d}y)\,\mathrm{d}s.\end{aligned}$$

We get therefore that for $f\in\mathcal{C}_c([0,T]\times\mathcal{K}_I)$, almost surely,

$$\int_0^T \int_{\mathcal{K}_I} f(s,x)\pi_s(\mathrm{d}x)\,\mathrm{d}s = \int_0^T \int_{\mathcal{K}_{I_{[2,p_2]}}} f\left(s,(y,\widetilde{L}_1(y))\right) \pi_s^{[2,p_2]}(\mathrm{d}y)\,\mathrm{d}s, \tag{4.48}$$

with the slight abuse of notation of writing $x=(x_{[2,p_2]},x_{[p_1]})$ for $x\in\mathbb{R}_+^{I^*}$.

We can now use our induction assumption to identify the measure $\mathrm{d}s\otimes\pi_s^{[2,p_2]}(\mathrm{d}y)$. To do so, we have to show that a set of equations as in Relation (4.40) holds for $\pi_s^{[2,p_2]}$ and an appropriate $\overline{\kappa}^2$.

If we can find some $\overline{\kappa}^2\in\Omega(I_{[2,p_2]})$ depending only on the initial κ such that for all $y\in\mathcal{K}_{I_{[2,p_2]}}$, for all $i\in I^*_{[2,p_2]}$,

$$\begin{aligned}\kappa_{0i} + \sum_{j\in I^*_{[p_1]}} \kappa_{ji}(L_i^1(y))^{p_1} + \sum_{j\in I^*_{[2,p_2]}\setminus\{i\}} & \kappa_{ji} y_j^{k_j} - \kappa_i^+ y_i^{k_i} \\ &= \overline{\kappa}^2_{0i} + \sum_{j\in I^*_{[2,p_2]}\setminus\{i\}} \overline{\kappa}^2_{ji} y_j^{k_j} - \overline{\kappa}_i^{2,+} y_i^{k_i}.\end{aligned} \tag{4.49}$$

Applying Relation (4.48) in Relation (4.40), for any $2\le p\le p_2$, for any $f\in\mathcal{C}^2\left(\mathcal{K}_{I_{[2,p]}}\right)$, almost surely, we have that the relation

$$\int_0^t \int_{\mathcal{K}_{I_{[2,p_2]}}} \sum_{i\in I^*_{[p]}} \left(\overline{\kappa}^2_{0i} + \sum_{j\in I^*\setminus\{i\}} \overline{\kappa}^2_{ji} x_j^{k_j} - \overline{\kappa}_i^{2,+} x_i^p\right) \frac{\partial f}{\partial x_i}(x_{[2,p]}) \pi_s^{[2,p_2]}(\mathrm{d}x)\,\mathrm{d}s = 0,$$

holds for $t\in[0,T]$. We recognize here the Relations of Proposition 4.14, for the set of indices $I_{[2,p_2]}$ and the matrix $\overline{\kappa}_2\in\Omega(I_{[2,p_2]})$. We can apply the induction hypothesis on the measure $\pi^{[2,p_2]}$. Setting $\widetilde{L}_2$ the unique solution of the system (4.28) of Proposition 4.9 for the set of indices $I_{[2,p_2]}$ and the matrix $\overline{\kappa}^2$, Relation (4.48) can be rewritten as : for $f\in\mathcal{C}_c([0,T]\times\mathcal{K}_I)$, almost surely,

$$\int_0^T \int_{\mathcal{K}_I} f(s,x)\pi_s(\mathrm{d}x)\,\mathrm{d}s = \int_0^T \int_{\mathcal{K}_{I_{[2,p_2]}}} f\left(s,(\widetilde{L}_2,\widetilde{L}_1(\widetilde{L}_2))\right)\mathrm{d}s, \tag{4.50}$$

with the slight abuse of notation of writing $x=(x_{[2,p_2]},x_{[p_1]})$ for $x\in\mathbb{R}_+^{I^*}$.

We conclude the induction by checking that

$$(\widetilde{L}_2,\widetilde{L}_1(\widetilde{L}_2)) = \ell_\kappa,$$

where $\ell_\kappa=(\ell_{\kappa,i})$ is the unique solution of the system (4.28) of Proposition 4.9.

For the existence of $\overline{\kappa}^2$ that verifies Relation (4.49). It is done by induction on the number of elements of the set $I^*_{[p_1]}$. If this set contains only one index i_0, setting $\overline{\kappa}^{i_0}$ such that for i, $j\in I_{[1,p_2]}$, $j\neq i$,

$$\overline{\kappa}^{i_0}_{ij} = \kappa_{ij} + \frac{\kappa_{ii_0}\kappa_{i_0j}}{\kappa^+_{i_0}}, \tag{4.51}$$

is suitable. Otherwise, if $I^*_{[p_1]}$ contains more than one element, we remove them, one by one, by applying the transformation of Relation (4.51).

The theorem is proved.

□

6. The General Case

We can now conclude the proof of Theorem 4.4. The difference with Section 5 is the time-inhomogeneity of the limiting quantities.

PROPOSITION 4.18. *If the initial conditions satisfy Relation* (4.15) *then the sequence of processes* $((X^N_{[1]}(t)), \Lambda_N)$, *defined by Relations* (4.14) *and* (4.16), *is tight for the convergence in distribution. Any limiting point* $((x(t)), \Lambda_\infty)$ *is such that*

(a) *Almost surely,* $(x(t))$ *is a continuous process with values in* $\mathcal{K}_{I_{[1]}}$;
(b) *For any function* $f \in \mathcal{C}_c([0,T]\times(\mathbb{R}_+)^{I^*})$,

$$\langle \Lambda_\infty, f\rangle = \int_{[0,T]\times\mathcal{K}_{I_{[2+]}}} f\,(s,(x(s),y))\,\pi^{[2+]}_s(\mathrm{d}y)\,\mathrm{d}s, \tag{4.52}$$

where $(\pi^{[2+]}_s)$ *is an optional process with values in* $\mathcal{P}(\mathcal{K}_{I_{[2+]}})$.

Recall the convention of writing an element x of $(\mathbb{R}_+)^{I^*}$ as $x=(x_{[1]}, x_{[2+]})$. See Section 2.3.

PROOF. The tightness of the occupation measures is shown exactly as in the proof of Proposition 4.12. Definition (4.27), Theorem 4.8 shows that the tightness of $(\overline{X}_N(t\wedge T_N))$ gives the tightness of the sequence of processes $(\overline{X}_N(t))$. It is established via the criterion of the modulus of continuity. See Theorem 7.3 of Billingsley [14].

For $i \in I^*_{[1]}$, $\delta>0$, Relation (4.10) gives the relation

$$w^N_i(\delta) \stackrel{\text{def.}}{=} \sup_{\substack{s,t\le T\wedge T_N\\ |s-t|\le\delta}} \left|\overline{X}^N_i(t)-\overline{X}^N_i(s)\right| \le \kappa_{0i}\delta+2\sup_{t\le T\wedge T_N}|\overline{M}_N(t)| + \sum_{j\in I^*\setminus\{i\}}\kappa_{ji}\int_s^t \overline{X}^N_j(u)^{(k_j)}\,\mathrm{d}u + \sum_{j\in I\setminus\{i\}}\kappa_{ij}\int_s^t \overline{X}^N_i(u)\,\mathrm{d}u,$$

where $(\overline{M}_N(t\wedge T_N))$ is a martingale whose previsible increasing process at time T is

$$\frac{k_i^2}{N}\left(\kappa_{0i}T\wedge T_N+\sum_{j\in I^*\setminus\{i\}}\kappa_{ji}\int_0^{T\wedge T_N}\overline{X}^N_j(u)^{(k_j)}\,\mathrm{d}u+\sum_{j\in I\setminus\{i\}}\kappa_{ij}\int_0^{T\wedge T_N}\overline{X}^N_i(u)^{(k_i)}\,\mathrm{d}u\right).$$

The expected value of this quantity on the event $\mathcal{E}_N$ converge to 0, by Doob's Inequality and Theorem 4.8, the martingale $(\overline{M}_N(t\wedge T_N))$ converges in distribution to 0. The proposition is proved. □

PROPOSITION 4.19. *If* $((x(t)), \Lambda_\infty)$ *is a limiting point of* $((\overline{X}_N(t)), \Lambda_N)$ *with the representation* (4.52), *then for* $p \ge 2$, *for* $f\in\mathcal{C}^2((\mathbb{R}^*_+)^{I^*_{[2,p]}})$, *almost surely, for*

all $t\in[0,T]$, the relation

$$\int_0^t \int_{\mathcal{K}_{I^*_{[2+]}}} \sum_{i\in I^*_{[p]}} \left(\sum_{j\in I^*_{[1]}} \kappa_{ji} x_j(s) + \kappa_{0i} + \sum_{j\in I^*_{[2+]}\setminus\{i\}} \kappa_{ji} y_j^{k_j} - \kappa_i^+ y_i^p \right) \frac{\partial f}{\partial x_i}(y_{[2,p]}) \pi_s^{[2+]}(\mathrm{d}y)\,\mathrm{d}s = 0. \tag{4.53}$$

holds.

Proof. We take a subsequence $((\overline{X}^{N_p}_{[1]}(t)), \Lambda_{N_p})$ converging in distribution to the random variable $((x(t)), \Lambda_\infty)$. The occupation measure of $(\overline{X}^N_i(t), i\in I_{[2+]})$ is converging in distribution to $\Lambda^{[2+]}_\infty$ defined by

$$\left\langle \Lambda^{[2+]}_\infty, f \right\rangle = \int_0^T \int_{\mathcal{K}_{I_{[2+]}}} g(y) \pi_s^{[2+]}(\mathrm{d}y)\,\mathrm{d}s,$$

for $f\in\mathcal{C}_c((\mathbb{R}_+)^{I_{[2+]}})$. Since the process $(\overline{X}^{N_p}_i(t), i\in I^*_{[1]})$ converges in distribution, for the uniform norm on $[0,T]$, we obtain a representation of Λ_∞,

$$\langle \Lambda_\infty, g\rangle = \int_0^T \int_{\mathcal{K}_I} g(y)\pi_s(\mathrm{d}y)\,\mathrm{d}s = \int_0^T \int_{\mathcal{K}_{I_{[2+]}}} g(x(s),y)\pi_s^{[2+]}(\mathrm{d}y)\,\mathrm{d}s, \tag{4.54}$$

for $g\in\mathcal{C}_c((\mathbb{R}_+)^{I^*})$. With the same method as in the proof of Proposition 4.14, the analogue of Relation (4.40) is established. We conclude the proof by using Relation (4.54). □

Proof of Theorem 4.4. In view of Theorem 4.17, we can assume $I_{[1]}\neq\emptyset$.

First, lets identify Λ_∞. Using Relation (4.54), we only have to identify the measure $\mathrm{d}s\otimes\pi_s^{[2+]}(\mathrm{d}y)$ on $\mathbb{R}_+\times(\mathbb{R}^*_+)^{I^*_{[2,+]}}$.

We set for $t\in[0,T]$, $\overline{\kappa}^3(t)$ as follows: For i, $j\in I^*_{[2+]}$, $j\neq i$, $\overline{\kappa}^3_{ij}(t)=\kappa_{ij}$ and

$$\overline{\kappa}^3_{0i}(t) = \kappa_{0i} + \sum_{j\in I^*_{[1]}} x_j(t)\kappa_{ji}, \quad \overline{\kappa}^3_{i0}(t) = \kappa_{i0} + \sum_{j\in I^*_{[1]}} \kappa_{ij}.$$

Relation (4.53), can be rewritten as Relation (4.40), for the set of indices $I_{[2+]}$ and the matrix $(\overline{\kappa}^3(t))\in\Omega(I_{[2,p_2]})^{[0,T]}$. The species $i\in I^*_{[1]}$ behave for the fast species as *exterior input*, with rate time dependent.

The convergence of the measure $\pi^{[2+]}(\mathrm{d}y)$ is then shown similarly as in the proof of Theorem 4.17. The only difference is the time dependence of the $\overline{\kappa}^3(t)$, which does not create any difficulty, since Proposition 4.18 gives the continuity of $(\overline{\kappa}^3(t))$ on $[0,T]$. Using Relation (4.54), for $g\in\mathcal{C}_c((\mathbb{R}_+)^{I^*})$, we have

$$\langle \Lambda_\infty, g\rangle = \int_0^T \int_{\mathcal{K}_I} g(y)\pi_s(\mathrm{d}y)\,\mathrm{d}s = \int_0^T g((x(s)),\ell(s))\,\mathrm{d}s,$$

where for all $t\in[0,T]$ $\ell(t)$ is the unique solution of the system (4.28) of Proposition 4.9 for the set of indices $I_{[2+]}$ and the matrix $\overline{\kappa}^3(t)$. It is easily seen that for all $t\in[0,T]$,

$$\ell(t) = L(x(t)),$$

where L is defined in Relation (4.19).

The convergence of the occupation measure is shown.

For the identification of the function $(x_{[1]}(t))=(x_i(t), i\in I^*_{[1]})$, integrating Relation (4.10) and dividing it by N, we get for $t\in[0,T]$, $i\in I^*_{[1]}$:

$$\overline{X}_i^N(t) = \overline{X}_i^N(0) + M_i^N(t) + \sum_{j\in I^*_{[1]}\setminus\{i\}} \int_0^t \kappa_{ji}\overline{X}_j^N(s)\,\mathrm{d}s + \int_0^t \sum_{j\in I_{[2+]}} \kappa_{ji}\frac{(X_j^N(s))^{(k_i)}}{N}\,\mathrm{d}s - \kappa_i^+\int_0^t \overline{X}_i^N(s)\,\mathrm{d}s, \tag{4.55}$$

where $(M_i^N(t))$ is a local martingale whose previsible increasing process is given by, for $t\le T$,

$$\left\langle M_i^N\right\rangle(t) = \frac{1}{N}\sum_{j\in I^*_{[1]}\setminus\{i\}} \int_0^t \kappa_{ji}\overline{X}_j^N(s)\,\mathrm{d}s + \frac{1}{N}\sum_{j\in I_{[2+]}} \int_0^t \kappa_{ji}\frac{(X_j^N(s))^{(k_i)}}{N}\,\mathrm{d}s + \frac{\kappa_i^+}{N}\int_0^t \overline{X}_i^N(s)\,\mathrm{d}s.$$

Using Doob's inequality and the bound of $(\overline{X}^N(t))$ on the event $\mathcal{E}_N$, we get the convergence in distribution of the martingales to 0.

Relation (4.42), Lemma 4.13, and the convergence of π_s^{2+} just proven, shows that for the convergence in distribution, for $j\in I^*_{[2+]}$,

$$\lim_{N\to+\infty}\left(\int_0^t \kappa_{ji}\frac{(X_j^N(s))^{(k_i)}}{N}\,\mathrm{d}s, t\in[0,T]\right) = \left(\int_0^t \kappa_{ji}(L_j(x(s)))^{k_j}\,\mathrm{d}s, t\in[0,T]\right),$$

and therefore, taking N to infinity in Relation (4.55), we get for $t\in[0,T]$, $i\in I^*_{[1]}$:

$$x_i(t) = \alpha_i + \sum_{j\in I^*_{[1]}\setminus\{i\}} \int_0^t \kappa_{ji}x_j(s)\,\mathrm{d}s + \int_0^t \sum_{j\in I_{[2+]}} \kappa_{ji}(L_j(x(s)))^{k_j}\,\mathrm{d}s - \kappa_i^+\int_0^t x_i(s)\,\mathrm{d}s,$$

which is exactly Relation (4.20).

Since $(x_{[1]}(t))$ lives in $\mathcal{K}_{I_{[1]}}$, the solution of this ODE is unique, and therefore the identification of $(x_{[1]}(t))$ is complete.

□

Note that ODE (4.20) can be rewritten as

$$\dot{x}_i(t) = \overline{\kappa}^4_{0i} + \sum_{j\in I^*_{[1]}\setminus\{i\}} x_j(t)\overline{\kappa}^4_{ji} - x_i(t)\sum_{j\in I\setminus\{i\}} \overline{\kappa}^4_{ij}, \quad i\in I^*_{[1]},$$

where $\overline{\kappa}^4\in\Omega(I_{[1]})$ is a matrix depending on the initial κ, constructed following the steps of the construction of $\overline{\kappa}^2$ in the Proof of Theorem 4.17. The $\overline{\kappa}^4$ can be given explicitly in terms of a path between complexes of $I^*_{[1]}$. The simplified ODE corresponds to the ODE associated to a CRN with only the complexes

$$\{\emptyset\}\cup\{S_i, i\in I^*_{[1]}\},$$

whith reactions defined by $\overline{\kappa}^4$. As an example, the limit $(x_4(t))$ of $(\overline{X}_4^N(t))$ in the CRN of Figure 1 is solution of the ODE associated to the CRN

$$\emptyset \underset{\overline{\kappa}^4_{40}}{\overset{\overline{\kappa}^4_{04}}{\rightleftharpoons}} S_4,$$

with

$$\overline{\kappa}^4_{04} = \frac{\kappa_{01}\kappa_{12}\kappa_{24}}{\kappa_1^+\kappa_2^+} + \frac{\kappa_{01}\kappa_{13}\kappa_{34}}{\kappa_1^+\kappa_3^+} + \frac{\kappa_{01}\kappa_{12}\kappa_{23}\kappa_{34}}{\kappa_1^+\kappa_2^+\kappa_3^+} \quad \text{and} \quad \overline{\kappa}^4_{40} = \frac{\kappa_{43}\kappa_{30}}{\kappa_3^+}.$$

CHAPTER 5

On the Recurrence Properties of Stochastic Chemical Reaction Networks with Two Species

Contents

1. Introduction

In this paper, we consider a general class of weakly reversible stochastic chemical reaction networks (CRNs) with two chemical species S_1 and S_2. They are referred to as 2D-CRNs, such a CRN can be described with a finite set of chemical reactions of the type

$$y_1^- S_1 + y_2^- S_2 \stackrel{\kappa}{\rightharpoonup} y_1^+ S_1 + y_2^+ S_2, \tag{5.1}$$

where $y^-=(y_1^-, y_2^-)$ and $y^+=(y_1^+, y_2^+)\in\mathbb{N}^2$.

With the kinetics of the *law of mass action*, see Voit et al. [76] and Lund [56], the time evolution of a 2D-CRN can be represented by a Markov process $(X(t))$ with values in $\mathbb{N}^2$. In state $x=(x_1, x_2)\in\mathbb{N}^2$, provided that $x_1\geq y_1^-$ and $x_2\geq y_2^-$, the transition associated to the reaction (5.1) is $x\to x+y^+-y^-$ and its rate is given

$$\kappa \frac{x_1!}{(x_1-y_1^-)!}\frac{x_2!}{(x_2-y_2^-)!} \sim \kappa x_1^{y_1^-} x_2^{y_2^-},$$

for x_1 and x_2 large.

The stability of this class of CRNs, i.e. the positive recurrence of $(X(t))$, is investigated in Agazzi et al. [2] by showing that the entropy function $(V(x))$ is a Lyapunov function for the infinitesimal generator of $(X(t))$, with, for $x=(x_1, x_2)\in\mathbb{R}_+^2$,

$$V(x) = v(x_1) + v(x_2) \quad \text{with} \quad v : y \in \mathbb{N} \mapsto y\ln(y) - y + 1, \tag{5.2}$$

with the convention $0\ln(0)=0$.

Our goal in this paper is of showing that, starting from a large state in the interior of $\mathbb{N}^2$, then there exists some $K>0$ such that

$$T_K \stackrel{\text{def.}}{=} \inf\{t\geq 0 : \min(X_1(t), X_2(t)) \leq K\}$$

is integrable and there exists a constant C_0 such that

$$\mathbb{E}_x(T_K) \leq C_0 V(x), \tag{5.3}$$

for all $x=(x_1, x_2)$ such that $\min(x_1, x_2)>K$.

This general result gives us an insight on the amount of time required for the process to return to the boundary of the domain. It is specific to the interior of the domain and uniform in the sense that the bound of Relation (5.3) with the function V holds for all such CRNs.

The proof is done with the use of Filonov's formulation of the Lyapunov condition introduced in Laurence and Robert [52] with the entropy function. One has to show that there exists positive constants γ, K and an integrable stopping time τ such that, for $x{=}(x_1,x_2){\in}\mathbb{N}^2$ with $\min(x_1,x_2){\geq}K$, then

$$\mathbb{E}_x(V(X(\tau)){-}V(X(0))) \leq -\gamma\mathbb{E}_x(\tau). \tag{5.4}$$

In this case, it can be shown that Relation (5.3) holds. Filonov's approach has already been used for specific examples of CRNs in Laurence and Robert [52] and Laurence and Robert [54] to prove the positive recurrence of the associated Markov processes. In this chapter we show that it can also be used in the context of a quite large class of CRNs.

To establish Relation (5.4), a finite partition of the interior of the state space is used to define a convenient stopping time τ. Ideas related to the notion of endotactic CRN developed in [10] are used. For some subsets of the partition, taking $\tau{=}t_1$ is enough, where t_1 is the first instant of jump of the process $(X(t))$. For the other cases, functional limit theorems for the Markov process on a convenient timescale are necessary to define the appropriate τ.

The approach of the stability of 2D-CRNs of Agazzi et al. [2] considers essentially the case when τ is taken as t_1. Finding γ and K such that Relation (5.4) holds for this choice of τ turns out to be quite technical in fact.

It should be noted that, starting from a neighborhood of the boundary, the time evolution of the process towards the origin is much more complicated. There does not exist a bound of the type (5.3) with a fixed function V independent of the CRN. See Section 7 of Laurence and Robert [52] for an example of 2D-CRN with a (very) slow return along one of the boundaries.

2. Model and main results

We introduce the main notations and definitions of this paper in this section. A general overview of the results is presented.

2.1. Chemical reaction network. In this paper, we study a general class of CRNs with two species (2D-CRN). Such a CRN is defined by a triple $\mathcal{X} = (\mathcal{S},\mathcal{C},\mathcal{R})$, where $\mathcal{S}=\{1,2\}$ is the set of species, reduced to two elements, the set of *complexes* $\mathcal{C}$ is a finite subset of $\mathbb{N}^2$, and the set of *reactions* $\mathcal{R}$ is a finite subset of $(\mathcal{C}\times\mathcal{C})\backslash\{(y,y), y\in\mathcal{C}\}$.

The species $i\in\mathcal{S}$ is often written S_i, and for a complex $y\in\mathcal{C}$, we use the notation

$$y = y_1S_1 + y_2S_2,$$

and $\emptyset$ refers to the complex associated to the null vector $(0,0)$. A reaction $(y^-,y^+){\in}\mathcal{R}$ is usually written $y^- \rightharpoonup y^+$.

The *reaction graph* associated to a CRN is the directed graph whose vertices are the complexes and whose set of directed edges is $\mathcal{R}$.

The connected components of the graph of the CRN determine a partition of the complexes into *linkage classes*. We set ℓ the number of linkage classes of the CRN studied, $(\mathcal{C}_i)$ the partition of $\mathcal{C}$, and $\mathcal{X}_i = (\mathcal{S},\mathcal{C}_i,\mathcal{R}_i)$ the CRN that compose the ith linkage class: for $1\leq i\leq\ell$,

$$\mathcal{R}_i \overset{\text{def.}}{=} \{(y^-,y^+)\in\mathcal{R} : y^-,y^+\in\mathcal{C}_i\}.$$

We make the assumption that the CRN studied is *weakly reversible* : each connected component of its reaction graph is strongly connected.

2.2. The Stochastic Markov process. The state of the CRN $\mathcal{X}$ is given by a vector $x{=}(x_1, x_2){\in}\mathbb{N}^2$, for $i = 1, 2$, x_i is the number of copies of chemical species S_i. A chemical reaction $r{=}y^- \rightharpoonup y^+$ corresponds to the change of state, for $x{=}(x_1, x_2)$,

$$x \longrightarrow x{+}y^+{-}y^-, \tag{5.5}$$

provided that $y_i^-{\leq}x_i$ holds for $i = 1, 2$, i.e. there are at least y_i^- copies of chemical species of type i, for all $i{\in}\mathcal{S}$, otherwise the reaction cannot happen.

The dynamical behavior of the system is governed by the *law of mass action*, see Voit et al. [76], Lund [56] for surveys on the law of mass action and the historical reference Guldberg and Waage [35]. A vector $\kappa{=}(\kappa_r, r{\in}\mathcal{R})$ of positive numbers is added to the parameters of the model. For $r \in \mathcal{R}$, κ_r is the *constant rate* of reaction r, and the rate of transition (5.5) in state $x \in \mathbb{N}^2$ is given by

$$\lambda_r(x) \stackrel{\text{def.}}{=} \kappa_r \frac{x_1!}{(x_1 - y_1^-)!} \frac{x_2!}{(x_2 - y_2^-)!} \mathbb{1}_{\{x_1 \geq y_1^-\}} \mathbb{1}_{\{x_2 \geq y_2^-\}}. \tag{5.6}$$

The process $(X(t))$ associated to the CRN $\mathcal{X}$ is defined by his infinitesimal generator $\mathcal{A}$:

$$\mathcal{A}(f)(x) = \sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} \lambda_r(x)(f(x + y^+ - y^-) - f(x)). \tag{5.7}$$

for any function $f : \mathbb{N}^2 \to \mathbb{R}$.

2.3. Filonov's Criterion. We will show that there exist an integrable stopping time τ and a constant $\gamma > 0$ such that Relation (5.4) holds for $x \in G$, for some convenient subset G of $\mathbb{N}^2$, where $(V(x))$ is the function defined by Relation (5.2).

When G^c is a finite subset of $\mathbb{N}^2$, Relation (5.4) is in fact sufficient to show the stability of the process $(X(t))$. This is Filonov's criterion. See Laurence and Robert [52].

We will show that Relation (5.4) holds *in the interior of the space of states*, away from the boundaries of the state space. Besides, since we consider general 2D-CRNs, the chemical species S_1 and S_2 have symmetrical roles, hence, considering states $x{=}(x_1, x_2){\in}\mathbb{N}^2$ such that $x_1 \geq x_2$ is enough. The convenient subset G is $B(K_1, K_2)$, with

$$B(K_1, K_2) \stackrel{\text{def.}}{=} \left\{x \in \mathbb{N}^2 : \|x\| \geq K_1 \quad \text{and} \quad x_1 \geq x_2 \geq K_2\right\}, \tag{5.8}$$

for $K_1 \geq K_2 \geq 1$, where $\|y\| = y_1 + y_2$, for $y \in \mathbb{N}^2$.

Such a subset is chosen so that the boundary behaviors of the system, mentioned in the introduction, do not impact our study. To ensure this, we will choose K_2 large enough so that if $x \in B(K_1, K_2)$, every reaction of the system can happen at state x:

$$K_2 \geq 2 \max\{y_i, \quad y \in \mathcal{C}, i = 1, 2\}.$$

When the stopping time τ in Relation (5.4) is chosen as the first instant of jump of the process, i.e. when $\tau = t_1$ with

$$t_1 \stackrel{\text{def.}}{=} \inf\{t \geq 0 : X(t-) \neq X(t)\}, \tag{5.9}$$

provided that convenient integrability properties hold, for $x \in \mathbb{N}^2$,

$$\mathbb{E}_x\left(V(X(t_1))\right) - V(x) = \mathbb{E}_x\left(\int_0^{t_1} \mathcal{A}(V)(x)\,\mathrm{d}s\right) = \mathcal{A}(V)(x)\mathbb{E}_x\left(t_1\right),$$

so that Relation (5.4) is equivalent to

$$\mathcal{A}(V)(x) \leq -\gamma. \tag{5.10}$$

It is usually this form of the criterion that is used in the literature of stochastic CRN. However, showing this inequality on $B(K_1, K_2)$ requires a complicated modification of the function V, defined piecewise on a partition of the space of states. For Filonov's criterion, we can choose to fix the entropy as the energy at all states, and then vary the definition of the stopping time τ to get the decrease of the energy.

2.4. Main result. The purpose of this paper is to show the following proposition :

THEOREM 5.1. *If $(X(t))$ is the Markov process associated to a weakly reversible 2D-CRN, irreducible on $\mathcal{E}_0 \subset \mathbb{N}^2$ infinite, there exists an integrable stopping time τ and constants $\gamma > 0$, $K_1 \geq K_2 \geq 1$ such that for all $x \in B(K_1, K_2)$,*

$$\mathbb{E}_x(V(X(\tau)) - V(X(0))) \leq -\gamma \mathbb{E}_x(\tau)$$

holds with V as the entropy function defined by Relation (5.2), and $B(K_1, K_2)$ is defined in Relation (5.8).

The following Corollary is a consequence of this theorem and Theorem 8.13 of Robert [67]. It shows that the process exits quite quickly the interior of the state space to hit a neighborhood of the boundary of the domain.

COROLLARY 5.2. *Let T_k the first time when $(X(t))$ gets to the boundary of width $k \geq 1$:*

$$T_k \stackrel{\text{def.}}{=} \inf\{t \geq 0 : \min\{X_1(t), X_2(t)\} \leq k\},$$

then there exist $K > 0$ such that

$$\mathbb{E}_x(T_K) \leq \frac{V(x)}{\gamma},$$

if $x = (x_1, x_2)$ is such that $\min(x_1, x_2) > K$.

This is a rather strong result, since the upper bound of the mean value of the stopping time does not depend closely on the CRN studied: only the constants K and γ depend on it. Starting from a state $x_n = (x_1^n, x_2^n)$ such that

$$\limsup_{n \to +\infty} \frac{\|x_n\|}{n} \leq 1,$$

we know that for some constant $K \geq 1$,

$$\limsup_{n \to +\infty} \frac{\mathbb{E}_{x_n}(T_K)}{n \ln(n)} < +\infty.$$

The strategy to use Filonov's criterion is to remark that once the energy of the system is fixed, here as the entropy function V, the stopping time can be chosen according to the initial state of the system. The space of states is partitioned into subsets on which a stopping time is defined. We will exploit Relation (5.10):

(a) At the states where we can show that Relation (5.10) holds, we will set the stopping time τ as t_1.
(b) For all the other states, a closer study of the time evolution of $(V(X(t)))$ will be necessary to set the right stopping time τ. Scaling arguments will be used.

2.5. Estimation of $\mathcal{A}(V)(x)$. We start by getting some estimate for $\mathcal{A}(V)(x)$, for $x \in B(K_1, K_2)$. To do so, we first estimate the contribution of one reaction $r = y^- \rightharpoonup y^+$ at state x, i.e. we estimate

$$\lambda_r(x)(V(x + y^+ - y^-) - V(x)),$$

for states of the form

$$x = x_N = (N, N^{\omega_2}),$$

where $\omega_2 \in (0, 1]$ and N is large. We set $\omega = (1, \omega_2)$.

At x_N, the following approximation holds :

$$\lambda_r(x_N) = \kappa_r \frac{x_1^N!}{(x_1^N - y_1^-)!} \frac{x_2^N!}{(x_2^N - y_2^-)!} = \kappa_r N^{y_1^- + \omega_2 y_2^-} + o(N^{y_1^- + \omega_2 y_2^-}).$$

Using the Taylor expansion of V, we get

$$V(x_N + y^+ - y^-) - V(x_N) = (y_1^+ - y_1^-)\ln(N) + (y_2^+ - y_2^-)\omega_2 \ln(N) + o(\ln(N)). \tag{5.11}$$

If $\langle \cdot, \cdot \rangle$ is the usual scalar product on $\mathbb{R}^2$, the following approximation holds:

$$\lambda_r(x_N)(V(x_N + y^+ - y^-) - V(x_N)) = \kappa_r \left\langle y^+ - y^-, \omega \right\rangle N^{\left\langle \omega, y^- \right\rangle} \ln(N) + o(N^{\left\langle \omega, y^- \right\rangle} \ln(N)).$$

As a consequence, if

$$\rho_0(\omega) \stackrel{\text{def.}}{=} \max\{\langle y, \omega \rangle,\ y \in \mathcal{C}\},$$

we can show that the dominant terms of $\mathcal{A}(V)(x_N)$ are the contributions of reactions $r = y^- \rightharpoonup y^+$, such that y^- verify $\langle y^-, \omega \rangle = \rho_0(\omega)$. In state x_N, these reactions are the *fastest reactions*.

This leads to the following estimate of $\mathcal{A}(V)(x_N)$:

$$\mathcal{A}(V)(x_N) = \sum_{\substack{r = y^- \rightharpoonup y^+ \in \mathcal{R}: \\ \left\langle y^-, \omega \right\rangle = \rho_0(\omega)}} \kappa_r (\left\langle y^+, \omega \right\rangle - \rho_0(\omega)) N^{\rho_0(\omega)} \ln(N) + o\left(N^{\rho_0(\omega)} \ln(N)\right). \tag{5.12}$$

Because of the maximality of $\rho_0(\omega)$, the terms $(\langle y^+, \omega \rangle - \rho_0(\omega))$ in the sum are negative or null. If one of these terms is non zero, this estimate is sufficient to show Relation (5.10) for N large enough. If not, one has to take the Taylor expansion to the next level.

The next step consist in understanding whether or not a reaction $r = y^- \rightharpoonup y^+$ such that $\langle y^-, \omega \rangle = \rho_0(\omega)$ and $\langle y^+ - y^-, \omega \rangle < 0$ exists.

2.6. Classification of the states. As mentioned above, to see the entropy decrease at state x_N, one has to find a dissipative reaction, i.e. a reaction $r = y^- \rightharpoonup y^+$ such that

$$\left\langle y^+ - y^-, \omega \right\rangle < 0.$$

If such a reaction exists, we set $\rho_d(\omega)$ as

$$\rho_d(\omega) \stackrel{\text{def.}}{=} \max\left\{\left\langle \omega, y^- \right\rangle : y^- \rightharpoonup y^+ \in \mathcal{R}, \left\langle y^+ - y^-, \omega \right\rangle < 0\right\}$$

Of the dissipative reactions, the ones with the largest rate verify $\langle y^-, \omega \rangle = \rho_d(\omega)$. If we have $\rho_d(\omega) = \rho_0(\omega)$, then Relation (5.12) allows us to conclude to Relation (5.10), and therefore to Relation (5.4) with $\tau = t_1$.

However, if $\rho_d(\omega) < \rho_0(\omega)$, Relation (5.12) is rewritten

$$\mathcal{A}(V)(x_N) = o\left(N^{\rho_0(\omega)} \ln(N)\right),$$

and we are not able to conclude yet.

More precisely, Relation (5.10) can be shown for each state in $B(K_1, K_2)$, except for the states in the sets $\mathcal{H}_c(\omega)$, for some $\omega_2 \in (0,1]$, $\omega = (1, \omega_2)$ and $c > 0$ such that

(a) ω is such that $\rho_d(\omega) < \rho_1(\omega)$
(b) The states in $\mathcal{H}_c(\omega)$ are precisely at the same order of magnitude as (N, N^{ω_2}). $\mathcal{H}_c(\omega)$ is defined as the set

$$\mathcal{H}_c(\omega) \stackrel{\text{def.}}{=} \left\{ x \in \mathbb{N}^2 : e^{-c} \leq \frac{x_2}{x_1^{\omega_2}} \leq e^c \right\}.$$

These states will be called H-states.

2.7. The H-states. Remains to show Relation (5.4) for the states in $\mathcal{H}_c(\omega)$, where ω verify

$$\rho_d(\omega) < \rho_0(\omega). \tag{5.13}$$

In this configuration, the fastest reactions (with rate $O(N^{\rho_0(\omega)})$) do not change the entropy of the system. One has to wait for a dissipative reaction, whose rate is $O(N^{\rho_d(\omega)})$, to happen.

In practice, we will use scaling arguments to show that on the timescale $t \mapsto t/N^{\rho_d(\omega)-1}$, the entropy of the system decreases significantly. We will show a result of the form, for $t_0 > 0$,

$$\limsup_{\|x\| \to +\infty,\ x \in \mathcal{H}_c(\omega)} \frac{\mathbb{E}_x\left(V(X(t_0/N^{\rho_d(\omega)-1}))\right)}{V(x)} < 1$$

from which Relation (5.4) will be deduced, setting $\tau = t_0/N^{\rho_d(\omega)-1}$.

2.8. Overview of the Paper. In Section 3 are introduced definitions and properties of the CRN used all through the proofs: we formally introduce the fastest reactions, the dissipative reactions, and characterize the ω such that Relation (5.13) holds.

We then proceed to the proof of Theorem 5.1. In order to prove Relation (5.10), we first study CRNs with one linkage class, in Section 4. In this case, at most one $\omega = (1, \omega_2)$ verifies Relation (5.13). Except in the set $\mathcal{H}_c(\omega)$ associated to this ω, we are able to show Relation (5.10).

In Section 5, using an argument of superposition for the infinitesimal generator, we deduce Relation (5.10) for the whole CRN, except for the H-states. Finally, in Section 6, we deal with the H-states using scaling arguments.

2.9. Notations. The following notations will be used throughout the paper. As a convention, when some $\omega_2 \in [0,1]$ is introduced, $\omega = (1, \omega_2)$.

For $z, z' \in \mathbb{R}^2$, we denote $\|z\| = |z_1| + |z_2|$ and the usual scalar product

$$\langle z, z' \rangle = z_1 z_1' + z_2 z_2'.$$

For $x \in \mathbb{R}_+^2$, $y \in \mathbb{R}_+^2$, we denote $x^y = x_1^{y_1} x_2^{y_2}$.

If $A \in \mathcal{B}(\mathbb{R}_+)$ is a Borelian subset of $\mathbb{R}_+$, and $\mathcal{P}$ is a Poisson process on $\mathbb{R}_+^2$, with intensity measure the Lebesgue measure on $\mathbb{R}_+^2$ we use the following notation,

$$\mathcal{P}(A, \mathrm{d}t) = \int_{x \in \mathbb{R}_+} \mathbb{1}_{\{x \in A\}} \mathcal{P}(\mathrm{d}x, \mathrm{d}t). \tag{5.14}$$

3. Properties of the 2D-CRN

Let $\mathcal{X} = \{\{S_1, S_2\}, \mathcal{C}, \mathcal{R}\}$ a weakly reversible 2D-CRN. In this section are introduces some definitions and properties associated to the graph of the CRN.

DEFINITION 5.3. *For some $\omega = (1, \omega_2) \in [0,1]^2$, we set*

$$\rho_0(\omega) \overset{\text{def.}}{=} \max\{\langle \omega, y\rangle\,,\ y \in \mathcal{C}\}, \tag{5.15}$$

and the set $I^1(\omega)$ as

$$I^1(\omega) \overset{\text{def.}}{=} \{y \in \mathcal{C} :\ \langle \omega, y\rangle = \rho_0(\omega)\}\,. \tag{5.16}$$

We also define the set

$$I^d(\omega) \overset{\text{def.}}{=} \{y \in \mathcal{C} :\quad \exists y \rightharpoonup y^+ \in \mathcal{R} : \left\langle y^+ - y, \omega\right\rangle < 0\} \tag{5.17}$$

and if $I^d(\omega) \neq \emptyset$,

$$\rho_d(\omega) = \max\{\langle \omega, y\rangle\,, y \in I^d(\omega)\}. \tag{5.18}$$

We use the convention $\rho_d(\omega) = -1$ if $I^d(\omega) = \emptyset$.

Note that since $\mathcal{X}$ is weakly reversible, each complex is source of at least one reaction.

We saw in Section 2 that for the states $x_N = (N, N^{\omega_2})$ for $\omega = (1, \omega_2) \in [0,1]^2$, and $N \gg 1$, the "classification" of the reactions and complexes can be done easily, using the sets of Definition 5.3.

- At state x_N, a reaction $y^- \rightharpoonup y^+$ has a rate proportional to $N^{\langle \omega, y^-\rangle}$, and changes the entropy of the order of $\langle y^+ - y^-, \omega\rangle \ln(N)$.
- The fastest reactions, i.e. the reactions with the largest rates, are the reactions starting from a complex in $I^1(\omega)$, and their rate is proportional to $N^{\rho_0(\omega)}$.
- The dissipative reactions, i.e. the reactions decreasing the entropy are reactions starting from a complex in $I^d(\omega)$. The fastest dissipative reactions have a rate proportional to $N^{\rho_d(\omega)}$.

The properties 5.4, 5.6, 5.8 and 5.9 are consequences of the weak reversibility of the CRN. They are mainly straightforward, they will be used repeatedly in the following. Their proof is given in the Appendix.

PROPOSITION 5.4. *The fastest reactions that change the entropy of the system are dissipative : for $\omega = (1, \omega_2) \in [0,1]^2$ and $y^- \rightharpoonup y^+ \in \mathcal{R}$,*

(a) *If $\langle \omega, y^-\rangle > \rho_d(\omega)$, then $\langle \omega, y^+ - y^-\rangle = 0$.*
(b) *If $\langle \omega, y^-\rangle = \rho_d(\omega)$, then $\langle \omega, y^+ - y^-\rangle \leq 0$.*
(c) *If $\langle \omega, y^+ - y^-\rangle > 0$, then $\langle \omega, y^-\rangle < \rho_d(\omega)$.*

Furthermore, if $I^d(\omega) \neq \emptyset$, then $I^1(\omega) \neq \mathcal{C}$.

In the next definition, we introduce the set of $\omega_2 \in [0,1]$ for which, at the states $x_N = (N, N^{\omega_2})$, the fastest reactions do not decrease the entropy.

DEFINITION 5.5. *Define $\mathcal{W}(\mathcal{X})$ as the set*

$$\mathcal{W}(\mathcal{X}) = \{\omega = (1, \omega_2) \in (0,1]^2 :\ I^1(\omega) \cap I^d(\omega) = \emptyset\}.$$

If $\omega \in \mathcal{W}(\mathcal{X})$, we define

$$\mathcal{H}_c(\omega) \overset{\text{def.}}{=} \left\{x \in \mathbb{N}^2 : e^{-c} \leq \frac{x_2}{x_1^{\omega_2}} \leq e^c\right\}.$$

for some $c > 0$. An element of $\mathcal{H}_c(\omega)$ is called an H-state.

Note that $(1,0) \notin \mathcal{W}(\mathcal{X})$.

At first, we will study weakly reversible 2D-CRNs with *one linkage class*. The following proposition gives the specific properties of such CRNs.

PROPOSITION 5.6. *Suppose $\mathcal{X}$ has only one linkage class. Let $\omega = (1, \omega_2) \in [0,1]^2$.*

a) *If $A \subset \mathcal{C}$ and $A \neq \mathcal{C}$, there are $y^- \in A$ and $y^+ \in \mathcal{C} \setminus A$ such that $y^- \rightharpoonup y^+ \in \mathcal{R}$.*
b) *If $I^d(\omega) \neq \emptyset$, then $I^1(\omega) \cap I^d(\omega) \neq \emptyset$.*
c) *If $I^1(\omega) \cap I^d(\omega) = \emptyset$, then for all $y \in \mathcal{C}$, $\langle y, \omega \rangle = \rho_0(\omega)$. Moreover, if $y \neq y'$, then $y_2 \neq y'_2$.*
d) *$\mathcal{W}(\mathcal{X})$ has at most one element.*

We are only interested in processes irreducible on an *infinite* subset of $\mathbb{N}^2$, hence the following definition.

DEFINITION 5.7. *The* stoichiometric compatibility class *associated to the CRN $\mathcal{X}$, from a state $x \in \mathbb{N}^2$, is the set $(x+S) \cap \mathbb{N}^2$, where*

$$S \stackrel{\text{def.}}{=} \left\{ \sum_{r=y^- \to y^+ \in \mathcal{R}} a_r (y^+ - y^-), \quad (a_r) \in \mathbb{Z}^{\mathcal{R}} \right\}.$$

The CRN $\mathcal{X}$ is said to have infinite stoichiometric compatibility classes *if the set $(x+S) \cap \mathbb{N}^2$ are infinite, except for a finite number of $x \in \mathbb{N}^2$.*

The following proposition characterize the dimension of the stoichiometric compatibility class according to the existence of a dissipative reaction.

PROPOSITION 5.8. *If $\mathcal{X}$ has infinite stoichiometric compatibility classes, $I^d(\omega) \neq \emptyset$ for all $\omega = (1, \omega_2) \in (0,1]^2$.*

This last proposition shows that their is a finite number of set of H-states:

PROPOSITION 5.9. *If the CRN $\mathcal{X}$ has ℓ linkage classes $(\mathcal{X}_i)$,*

$$\mathcal{W}(\mathcal{X}) \subset \cup_{i=1}^{\ell} \mathcal{W}(\mathcal{X}_i),$$

and $\mathcal{W}(\mathcal{X})$ has at most ℓ elements.

4. 2D-CRN with one linkage class

In this section, we study a 2D-CRN $\mathcal{X}$ with *one* linkage class.

PROPOSITION 5.10. *If $\mathcal{X}$ has one linkage class ($\ell = 1$), then*

(i) *If $\mathcal{W}(\mathcal{X}) = \emptyset$, we can find some $K_1 \geq K_2 \geq 1$ such that for all $x \in B(K_1, K_2)$, the inequality $\mathcal{A}(V)(x) \leq -1$ holds,*
(ii) *If $\mathcal{W}(\mathcal{X}) = \{\omega\}$, we can find some $K_1 \geq K_2 \geq 1$ and some $c > 0$ such that for all $x \in B(K_1, K_2) \setminus \mathcal{H}_c(\omega)$, the inequality $\mathcal{A}(V)(x) \leq -1$ holds,*

where $\mathcal{A}$ is the infinitesimal generator of the process associated to $\mathcal{X}$, V is the entropy function defined in Relation (5.2), B is defined in Relation (5.8) and $\mathcal{W}$ and $\mathcal{H}_c(\omega)$ in Definition 5.5.

The proof of Proposition 5.10 is carried out by contradiction: in (ii), if $\omega \in \mathcal{W}(\mathcal{X})$, we make the assumption that a sequence (x_p) with value in $B(K_1, K_2) \setminus \mathcal{H}_c(\omega)$ verifies $\|x_p\| \to +\infty$ and $\mathcal{A}(V)(x_p) > -1$ for all p.

The goal is then to show that for some subsequence (x_{p_n}), we have $\mathcal{A}(V)(x_{p_n}) \to -\infty$ which contradicts the former hypothesis.

To do so, we distinguish two situations, according to the limit of the sequence $\varepsilon_p = \ln(x_{p,2}) / \ln(x_{p,1})$:

(a) If the sequence is converging to the boundary of $\mathcal{H}_c(\omega)$, i.e. if $\varepsilon_p \to \omega_2$, we are able to give an approximation of $\mathcal{A}(V)(x_p)$ at an order $O(x_{p,1}^{\rho_0(\omega)})$ (one step further than the approximation given in Relation (5.12)), which goes to $-\infty$ when $x_{p,1}$ goes to $+\infty$.
(b) If (ε_p) converges to $\varepsilon \neq \omega_2$, with similar ideas as in Anderson and Kim [8]: we classify the rates of the reactions *along the sequence* (x_p), and show that at least one of the fastest reactions along (x_p) is dissipative.

Proof of Proposition 5.10. The following simple inequalities are used throughout the paper: For $x > 1$, for $a \in \mathbb{R}$, $2|a| < x$, for the function $v : x \mapsto x\ln(x) - x + 1$,

$$|v(x+a) - v(x)| \leq 2|a|\ln(x) \quad \text{and} \quad |v(x+a) - v(x) - a\ln(x)| \leq \frac{|a|^2}{x-|a|} \leq \frac{2|a|^2}{x}. \tag{5.19}$$

They can be shown using Taylor's theorem.

This proof is done in the case where $\omega = (1, \omega_2) \in [0,1]^2$ is such that $I^1(\omega) \cap I^d(\omega) = \emptyset$, i.e. in situation (i) of the proposition, or in situation (ii) with $\omega = (1,0)$. It can be easily adapted to the case where no such ω exists.

First, recall that we assume

$$K_2 \geq 2\max\{y_i,\ y \in \mathcal{C}, i = 1,2\}$$

so that at $x \in B(K_1, K_2)$, each reaction is "active" (i.e. $\lambda_r(x) > 0$ for each reaction).

Let $c > 0$, $K_1 \geq K_2 \geq 1$ specified later, and $(x_p) \in B(K_1, K_2) \setminus \mathcal{H}_c(\omega)$ a sequence of states such that

$$\lim_{p \to +\infty} \|x_p\| = +\infty$$

and by contradiction, for all $p \geq 1$,

$$\mathcal{A}(V)(x_p) > -1. \tag{5.20}$$

Until the end of the proof, we may take several subsequences of the sequence (x_p), that we will still write (x_p).

We set the sequence

$$\varepsilon_p = \frac{\ln(x_{p,2})}{\ln(x_{p,1})}.$$

By definition, $\varepsilon_p \in [0,1]$ for all $p \geq 1$. Up to a subsequence, we can assume that (ε_p) converges to $\varepsilon \in [0,1]$.

1) **If $\varepsilon = \omega_2$.**

Since $x_p \notin \mathcal{H}_c(\omega)$,

$$\nu_p \overset{\text{def.}}{=} (\varepsilon_p - \omega_2)\ln(x_{p,1}) \geq c \quad \text{or} \quad \nu_p \leq -c.$$

Note that if $\omega_2 = 0$, resp. $\omega_2 = 1$, then necessarily, $\nu_p \geq c$, resp. $\nu_p \leq -c$.

Up to a sub-sequence, we can make the assumption that one of the inequalities is always valid. Here, lets assume that $\nu_p \geq c$ for all p, the other side is dealt in the same way.

For all $y \in \mathcal{C}$, $\langle y, \omega \rangle = \rho_0(\omega)$, see Proposition 5.6. Therefore, for any reaction $r = y^- \rightharpoonup y^+$,

$$\left\langle y^+ - y^-, (1, \varepsilon_p) \right\rangle \ln(x_{p,1}) = (y_2^+ - y_2^-)\nu_p,$$

and by Relation (5.19),

$$\frac{\lambda_r(x_p)(V(x_p + y^+ - y^-) - V(x))}{x_{p,1}^{\rho_0(\omega)}} \leq \kappa_r e^{\nu_p y_2^-} \left[(y_2^+ - y_2^-)\nu_p + \frac{2\|y^+ - y^-\|^2}{K_2} \right].$$

and therefore $\mathcal{A}(V)(x_p) \leq x_{p,1}^{\rho_0(\omega)} g(\nu_p)$ where g is as follows, for $u \in \mathbb{R}$,

$$g(u) \stackrel{\text{def.}}{=} \sum_{r=y^- \rightharpoonup y^+} \kappa_r e^{uy_2^-} \left[(y_2^+ - y_2^-)u + \frac{2\|y^+ - y^-\|^2}{K_2} \right].$$

Setting $k_m \stackrel{\text{def.}}{=} \max\{y_2^-, y \in \mathcal{C}\} = \max\{y_2, y \in \mathcal{C}\}$ and $y_m = (y_{1,m}, k_m)$ the (unique) associated complex, one gets

$$\lim_{u \to +\infty} \frac{1}{g(u)} \left(\sum_{r=y_m \rightharpoonup y^+ \in \mathcal{R}} \kappa_r e^{yk_m} (y_2^+ - k_m)u \right) = 1.$$

In the last sum, $y_2^+ - k_m < 0$ holds for each term because of Proposition 5.6 c) and therefore all terms in the sum are negative. Hence, we can find some $u_0 > 0$ such that $g(u) < -1$ for all $u > u_0$.

We set $c > u_0$ in the definition of $\mathcal{H}_c(\omega)$, and if $\omega = (1, 0)$, we set $K_2 \geq e^c$.

This choice of constant contradicts Condition (5.20) and therefore we can't have $\varepsilon = \omega_2$.

2) **If $\varepsilon \neq \omega_2$.**

We order the reactions "along the sequence (x_p)". Applying Lemma 5.1 of Anderson et al. [9], one can construct a subsequence of (x_p), and define a set $T^1_{(x_p)} \subset \mathcal{C}$ such that the following limits hold :

— For $y \in T^1_{(x_p)}$ and $y' \in \mathcal{C}$,

$$\lim_{p \to +\infty} \frac{x_p^{y'}}{x_p^y} < +\infty.$$

— For $y, y' \in T^1_{(x_p)}$, there exists some $C > 0$ such that

$$\lim_{p \to +\infty} \frac{x_p^y}{x_p^{y'}} = C.$$

— For $y \in T^1_{(x_p)}$, if $y' \notin T^1_{(x_p)}$,

$$\lim_{p \to +\infty} \frac{x_p^{y'}}{x_p^y} = 0.$$

The construction of the subsequence can be found in the proof of Lemma 4.2 in Anderson [5], and relies mainly in taking a subsequence for which the following order holds, for some indexation of the complexes $\{y_k, 1 \leq k \leq |\mathcal{C}|\} = \mathcal{C}$:

$$x_p^{y_1} \geq x_p^{y_2} \geq \ldots \geq x_p^{y_{|\mathcal{C}|}}.$$

The reactions starting from the complexes in $T^1_{(x_p)}$ are the *fastest reactions along* (x_p), meaning that when p goes to infinity, in state x_p, these reactions have the largest rate.

We now estimate the value of $\mathcal{A}(V)(x_p)$. Lets set for $x \in \mathbb{N}^2$,

$$D_1(x) \overset{\text{def.}}{=} \sum_{\substack{r=y^- \rightharpoonup y^+ \in \mathcal{R} \\ y^- \notin T^1_{(x_p)}}} \lambda_r(x)(V(x+y^+-y^-)-V(x)),$$

$$D_2(x) \overset{\text{def.}}{=} \sum_{\substack{r=y^- \rightharpoonup y^+ \in \mathcal{R} \\ y^-, y^+ \in T^1_{(x_p)}}} \lambda_r(x)(V(x+y^+-y^-)-V(x)),$$

$$D_3(x) \overset{\text{def.}}{=} \sum_{\substack{r=y^- \rightharpoonup y^+ \in \mathcal{R} \\ y^- \in T^1_{(x_p)}, y^+ \notin T^1_{(x_p)}}} \lambda_r(x)(V(x+y^+-y^-)-V(x)).$$

For all x, $\mathcal{A}(V)(x) = D_1(x) + D_2(x) + D_3(x)$. We deal with each part separately.

First note that for a reaction $r = y^- \rightharpoonup y^+$, for some $0 < c_r < C_r$,

$$c_r \leq \lim_{p \to +\infty} \frac{\lambda_r(x_p)}{x_p^{y^-}} < C_r.$$

Let $y_T \in T^1_{(x_p)}$.

— $D_1(x)$ is the contribution of the slow reactions along (x_p). For a reaction $r = y^- \rightharpoonup y^+$ such that $y^- \notin T^1_{(x_p)}$, we have

$$\limsup_{p \to +\infty} \frac{\lambda_r(x_p)}{x_p^{y^-}} \, \frac{x_p^{y^-}}{x_p^{y_T}} \, \ln\left(\frac{x_p^{y_T}}{x_p^{y^-}}\right) = 0,$$

and

$$\limsup_{p \to +\infty} \frac{\lambda_r(x_p)}{x_p^{y_T}} \ln\left(\frac{x_p^{y^+}}{x_p^{y_T}}\right) = 0.$$

Using Relation (5.19), this leads to

$$\lim_{p \to +\infty} \frac{D_1(x_p)}{x_p^{y_T}} = 0.$$

— $D_2(x)$ is the contribution of the fast reactions along (x_p), that do not decrease the entropy. For a reaction $r = y^- \rightharpoonup y^+ \in \mathcal{R}$ such that $y^-, y^+ \in T^1_{(x_p)}$, we have

$$\limsup_{p \to +\infty} \frac{\lambda_r(x_p)}{x_p^{y_T}} \leq C_2,$$

for some constant C_2, and using Relation (5.19),

$$|V(x_p + y^+ - y^-) - V(x_p)| \leq \left|\ln\left(\frac{x_p^{y^+}}{x_p^{y^-}}\right)\right| + C_3 \leq C_4.$$

for some constant $0 < C_3 < C_4$. Therefore,

$$\limsup_{p \to +\infty} \frac{D_2(x_p)}{x_p^{y_T}} \leq C_5,$$

for some $C_5 > 0$.

— $D_3(x)$ is the contribution of the fast reactions along (x_p), that decrease the entropy.

First, lets show that this sum is not empty. Let $\omega' = (1, \varepsilon)$. Since $\omega' \neq \omega$, $I^d(\omega') \neq \emptyset$, and therefore $I^1(\omega') \neq \mathcal{C}$, see Proposition 5.4. Hence, using Proposition 5.6 a), showing that $T^1_{(x_p)} \subset I^1(\omega')$ is sufficient.

For $y \notin I^1(\omega')$, and $y_0 \in I^1(\omega')$, $\langle y - y_0, \omega' \rangle = -\eta < 0$ and for p large enough, $|(y_2 - y_{0,2})(\varepsilon_p - \varepsilon)| < \eta/2$. Therefore,

$$\lim_{p\to+\infty} \frac{x_p^y}{x_p^{y_0}} = \lim_{p\to+\infty} x_{p,1}^{\langle y'-y_0,\omega'\rangle+(\varepsilon_p-\varepsilon)(y_2-y_{0,2})} = 0,$$

and $T^1_{(x_p)} \subset I^1(\omega')$.
We now estimate this sum. For $r = y^- \rightharpoonup y^+ \in \mathcal{R}$ such that $y^- \in T^1_{(x_p)}$, $y^+ \notin T^1_{(x_p)}$, using Relation (5.19), one gets

$$V(x_p + y^+ - y^-) - V(x_p)) \leq \ln\left(\frac{x_p^{y^+}}{x_p^{y^-}}\right) + C_6 \to -\infty,$$

and since besides,

$$\lim_{p\to+\infty} \frac{\lambda_r(x_p)}{x_p^{y_T}} = C_r > 0,$$

one gets

$$\lim_{p\to+\infty} \frac{D_3(x_p)}{x_p^{y_T}} = -\infty.$$

Summing up all these estimates, we get

$$\lim_{p\to+\infty} \mathcal{A}(V)(x_p) = -\infty,$$

which leads once again to a contradiction of Condition (5.20).
The proposition is proved. □

5. General 2D-CRN : Superposition of the linkage classes

We now study a general weakly reversible 2D-CRN $\mathcal{X}$, with ℓ linkage classes $(\mathcal{X}_i, 1 \leq i \leq \ell)$, for $\ell \geq 1$. We set $(X(t))$ its associated Markov chain, and $\mathcal{A}$ its infinitesimal generator. To use the results of Section 4, we use the following superposition relation :

$$\mathcal{A}(f)(x) = \sum_{i=1}^{\ell} \mathcal{A}_i(f)(x) \tag{5.21}$$
$$\text{with} \quad \mathcal{A}_i(f)(x) = \sum_{r=(y^-,y^+)\in\mathcal{R}_i} \lambda_r(x)(f(x + y^+ - y^-) - f(x)).$$

5.1. An example. Lets sum up our advances on the following example:

$$2S_1 \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} S_1 + 2S_2, \qquad 3S_2 \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} S_1. \tag{5.22}$$

This CRN has two linkage classes, $\mathcal{C}_1 = \{2S_1, S_1+S_2\}$ and $\mathcal{C}_2 = \{3S_2, S_1\}$. It is easy to verify that in this case,

$$\mathcal{W}(\mathcal{X}) = \mathcal{W}(\mathcal{X}_1) = \{(1, 1/2)\} \quad \text{and} \quad \mathcal{W}(\mathcal{X}_2) = \{(1, 1/3)\}.$$

Therefore, as a consequence of Proposition 5.10, we can find some $c_1, c_2 > 0$, $K_1 \geq K_2 \geq 1$ such that for $x \in B(K_1, K_2)$, $x \notin \mathcal{H}_{c_1}((1, 1/2)) \cup \mathcal{H}_{c_2}((1, 1/3))$,

$$\mathcal{A}(V)(x) = \mathcal{A}_1(V)(x) + \mathcal{A}_2(V)(x) \leq -2.$$

We still have to deal with the H-states of both CRN $\mathcal{X}_1$ and $\mathcal{X}_2$, i.e. with both sets $\mathcal{H}_{c_1}((1, 1/2))$ and $\mathcal{H}_{c_2}((1, 1/3))$.

In this section, it is the set $\mathcal{H}_{c_2}((1,1/3))$ that is studied. Even if H-states for the CRN $\mathcal{X}_2$, for the whole CRN, these states are not H-states since $(1,1/3) \notin \mathcal{W}(\mathcal{X})$.

The states in this set are of order (N, N^{ω_2}) with $\omega_2 = 1/3$, and in these states, the fastest reaction is $2S_1 \rightharpoonup S_1 + 2S_2$, which is dissipative :

$$\langle (S_1 + 2S_2) - 2S_1, \omega \rangle = -1/3 < 0.$$

Therefore, we are in the simple situation described in Section 2.5: we know precisely the order of magnitude of the state, which will enable us to show a similar Relation as (5.12).

5.2. The false H-states.

PROPOSITION 5.11. *We can find some $K_1 \geq K_2 \geq 1$, and some $c_\omega > 0$ for each $\omega \in \mathcal{W}(\mathcal{X}))$ such that, for all*

$$x \in B(K_1, K_2) \setminus \left(\bigcup_{\omega \in \mathcal{W}(\mathcal{X})} \mathcal{H}_{c_\omega}(\omega) \right),$$

the inequality $\mathcal{A}(V)(x) \leq -1$ holds.

Using Relation (5.21), we know that in all the states x such that $\mathcal{A}_i(V)(x) \leq -1$ for all $1 \leq i \leq \ell$, $\mathcal{A}(V)(x) \leq -1$ holds. To conclude with the result, we only have to prove this Relation for the states in $\mathcal{H}_c(\omega)$ for the ω such that $\omega \in \mathcal{W}(\mathcal{X}_i)$ for some i, but $\omega \notin \mathcal{W}(\mathcal{X})$. This is done by contradiction, similarly as in the proof of Proposition 5.10.

PROOF. We set

$$\overline{\mathcal{W}} \stackrel{\text{def.}}{=} \bigcup_{i=1}^{\ell} \mathcal{W}(\mathcal{X}_i).$$

As a consequence of Proposition 5.10 and Relation (5.21), the inequality $\mathcal{A}(V)(x) \leq -1$ is verified for x in

$$B(\overline{K}_1, \overline{K}_2) \setminus \left(\bigcup_{\omega \in \overline{\mathcal{W}}} \mathcal{H}^{\bar{c}_\omega}(\omega) \right),$$

for some $\overline{K}_1 \geq \overline{K}_2 \geq 1$ and $\bar{c}_\omega > 0$.

If $\omega \in \overline{\mathcal{W}}$ but $\omega \notin \mathcal{W}(\mathcal{X})$, lets show that we can find some $K_1^\omega \geq \overline{K}_1$, $K_2^\omega \geq \overline{K}_2$ such that, for $x \in B(K_1^\omega, K_2^\omega) \cap \mathcal{H}^{\bar{c}_\omega}(\omega)$, the inequality $\mathcal{A}(V)(x) \leq -1$ holds.

The proof is identical to the second part (when $\varepsilon \neq \omega_2$) of the Proof of Proposition 5.10. By contradiction, we take (x_p) a sequence of states of $\mathcal{H}(\omega)^{\bar{c}_\omega}(\omega)$ that verify $\mathcal{A}(V)(x_p) > -1$ for all $p \geq 1$, and such that

$$\lim_{p \to +\infty} \|x_p\| = +\infty.$$

The fact that $x_p \in \mathcal{H}^{\bar{c}_\omega}(\omega)$ allows an easier partition of the different complexes according to the speed of the reactions. Indeed, we can show that in this situation, $T^1_{(x_p)} = I^1(\omega)$, and that

$$\lim_{p \to +\infty} \frac{\mathcal{A}(V)(x_p)}{x_{p,1}^{\rho_0(\omega)}} = -\infty,$$

which contradicts the first hypothesis, and allows to conclude the proof.

□

6. A scaling argument for the H-states

In this section, we find a stopping time τ for the last states still remaining: the H-states. The 2D-CRN is taken weakly reversible, with an infinite stoichiometric compatibility classes, see Definition 5.7.

PROPOSITION 5.12. *If $\omega \in \mathcal{W}(\mathcal{X})$, $\omega_2 < 1$ and $c > 0$, one can find some stopping time τ, and some $0 < b < 1$ such that*

$$\limsup_{N\to+\infty} \sup_{(N,x_2^N)\in\mathcal{H}_c(\omega)} \frac{\mathbb{E}_{(N,x_2^N)}(V(X(\tau)))}{N\ln(N)} \leq b$$

$$\text{and} \quad \limsup_{N\to+\infty} \sup_{(N,x_2^N)\in\mathcal{H}_c(\omega)} \frac{\mathbb{E}_{(N,x_2^N)}(\tau)}{N} = 0.$$

6.1. Ideas behind the proof. In this section, we focus on the states $x_N = (N, x_2^N)$ such that

$$e^{-c}N^{\omega_2} \leq x_2^N \leq e^c N^{\omega_2},$$

and we call $(X^N(t))$ the process starting at x_N. Since $\omega \in \mathcal{W}(\mathcal{X})$, the fastest reactions at this state are not dissipative, and one has to wait for a reaction with smaller rate to occur to see the energy decrease. This phenomenon is well described by looking at the *different timescales* present in the system.

Let us sum up the role of each reaction here:

(a) The fastest reactions are the reactions $y^- \rightharpoonup y^+$ such that $y^- \in I^1(\omega)$. Here, all these reactions verify $\langle \omega, y^+ - y^- \rangle = 0$, and *do not dissipate* the entropy of the system. Therefore, starting from x_N, since these are the reactions that happen first with high probability, we can't choose $\tau = t_1$ in Relation (5.4)

(b) The reactions $y^- \rightharpoonup y^+$ that dissipate the entropy of the system verify

$$\langle y^+ - y^-, \omega \rangle < 0.$$

These are the reactions that we want to see happening. In particular, the fastest of these reactions will give us the "right timescale" to see the decrease of the entropy of the system. The reactions we are interested in are the reactions $y^- \rightharpoonup y^+$ that are dissipative and such that $\langle y^-, \omega \rangle = \rho_d(\omega)$, where $\rho_d(\omega)$ is introduced in Definition 5.3.

To study the evolution of $(V(X^N(t)))$, making the assumption that $(X_2^N(t))$ remains $O(N^{\omega_2})$ on the studied timescale, we use the following approximation :

$$\frac{V(X^N(t))}{N\ln(N)} = \frac{X_1^N(t)}{N} + o(1) = \frac{\langle X^N(t), \omega \rangle}{N} + o(1).$$

This motivates the study of the process $(S^N(t)) = (\langle X^N(t), \omega \rangle)$.

This process has the advantage of not being disturbed by the fast reactions. In fact, the fastest reactions changing $(S^N(t))$ are the reactions $r = y^- \rightharpoonup y^+$ such that $\langle y^-, \omega \rangle = \rho_d(\omega)$, and because of Proposition 5.4, they verify $\langle y^+ - y^-, \omega \rangle < 0$, and therefore are *decreasing* the process $(S^N(t))$.

Since we need to see $(S^N(t)/N)$ decrease significantly, we need to see $O(N)$ such reactions to happen, and therefore the relevant timescale here is

$$t \mapsto \frac{t}{N^{\rho_d(\omega)-1}}.$$

The rest of the proof uses standard scaling methods.

The proof is carried out in three steps:

— We first show that $(X_2^N(t/N^{\rho_d(\omega)-1}))$ stays of order N^{ω_2}, using a coupling argument and the upper bound of an $M/M/1$ queue, see Proposition 5.15 of Appendix 7.
— We then show that $(S^N(t/N^{\rho_d(\omega)-1})/N)$ can't increase.
— Finally, we show using a scaling argument that $(S^N(t/N^{\rho_d(\omega)-1})/N)$ has decreased significantly after a time $t_0 > 0$.

6.2. Formal proof.

Proof. Since the process is weakly reversible and has infinite stoichiometric compatibility classes, Proposition 5.8 shows that $I^d(\omega) \neq \emptyset$. We introduce $\rho_0(\omega)$ and $\rho_d(\omega)$ as in Definition 5.3, and $\gamma = \rho_d(\omega) - 1 > -1$. Recall that $\omega_2 > 0$.

For $N \geq 1$, let $x_N = (N, x_2^N) \in \mathcal{H}_c(\omega)$. x_2^N verifies the Relation

$$e^{-c} \leq \frac{x_2^N}{N^{\omega_2}} \leq e^c. \tag{5.23}$$

$(X^N(t))$ is the process associated to $\mathcal{X}$ starting at x_N, and solution of the following stochastic differential equation (SDE). See Section 2 of Laurence and Robert [52].

$$\mathrm{d}X^N(t) = \sum_{r=y^- \rightharpoonup y^+ \in \mathcal{R}} (y^+ - y^-)\mathcal{P}_r\left(\left(0, \kappa_r(X_N(t-))^{(y^-)}\right), \mathrm{d}t\right), \tag{5.24}$$

where $(\mathcal{P}_r, r \in \mathcal{R})$ is a family of independent Poisson processes on $\mathbb{R}_+^2$, with intensity measure the Lebesgue measure on $\mathbb{R}_+^2$, and for $z \in \mathbb{N}^2$, $y \in \mathbb{N}^2$,

$$z^{(y)} = \frac{z_1!}{(z_1 - y_1)!}\frac{z_2!}{(z_2 - y_2)!}\mathbb{1}_{\{z_1 \geq y_1\}}\mathbb{1}_{\{z_2 \geq y_2\}}.$$

We set $S^N(t) = \left\langle X^N(t), \omega \right\rangle$, and for all $N \geq 1$,

$$(\overline{S}^N(t), \overline{X}_1^N(t), \overline{X}_2^N(t)) \stackrel{\text{def.}}{=} \left(\frac{S^N(t/N^\gamma)}{N}, \frac{X_1^N(t/N^\gamma)}{N}, \frac{X_2^N(t/N^\gamma)}{N^{\omega_2}}\right)$$

Lets set

$$\begin{aligned}
H_{1+}^N &\stackrel{\text{def.}}{=} \inf\left\{t \geq 0 : \overline{X}_1^N(t) \geq M_1\right\},\\
H_{1-}^N &\stackrel{\text{def.}}{=} \inf\left\{t \geq 0 : \overline{X}_1^N(t) \leq m_1\right\},\\
H_{2+}^N &\stackrel{\text{def.}}{=} \inf\left\{t \geq 0 : \overline{X}_2^N(t) \geq M_2^{\omega_2}\right\},\\
H_{2-}^N &\stackrel{\text{def.}}{=} \inf\left\{t \geq 0 : \overline{X}_2^N(t) \leq m_2^{\omega_2}\right\},
\end{aligned}$$

and $H_i^N \stackrel{\text{def.}}{=} \min\{H_{i+}^N, H_{i-}^N\}$ for $i = 1, 2$, for some constant $m_1 < 1 < M_1$ and $0 < m_2 < e^{-c/\omega_2}/2 < 2e^{c/\omega_2} < M_2$ specified later.

Let $t_0 > 0$, to be specified later, and set

$$\overline{\tau}_N = \min\{t_0, H_1^N, H_2^N\} \quad \text{and} \quad \tau_N = \frac{\overline{\tau}_N}{N^\gamma}. \tag{5.25}$$

We first show that

$$\limsup_{N \to +\infty} \sup_{(N, x_2^N) \in \mathcal{H}_c(\omega)} \mathbb{E}_{(N, x_2^N)}(\overline{S}^N(\overline{\tau}_N)) \leq b \tag{5.26}$$

holds for some $0 < b < 1$, and some choice of $m_2, M_2, t_0 > 0$.

Control of H_2^N: Uniform bound of $(\overline{X}_2^N(t))$. We show that

$$\lim_{N\to+\infty} \mathbb{P}(H_2^N < t_0 \wedge H_1^N) = 0.$$

The bound is shown in two steps, the upper and the lower-bound. Since both steps are very similar, we will only do here the proof for the upper-bound, i.e. show that $\lim_{N\to+\infty} \mathbb{P}(H_{2+}^N < t_0 \wedge H_1^N) = 0$. The process $(X_2^N(t))$ is mainly governed by the *fastest reactions*, i.e. the reactions starting from a complex in $I^1(\omega)$. For the sake of simplicity, in this proof we don't take into account the other reactions. It is not very hard to generalize it to the real process. Similarly, we will use the following approximation $\lambda_r(x) \approx \kappa_r x^{y^-}$ for the rates of reaction $r = y^- \rightharpoonup y^+$, which is valid since both species are in large number.

Let $k_m = \max\{y_2, y \in I^1(\omega)\}$, and $r_* = y_*^- \rightharpoonup y_*^+$ a reaction such that $y_*^- \in I^1(\omega)$ and $y_{*,2}^- = k_m$. Note that since $\omega \in \mathcal{W}(\mathcal{X})$,

— if $y^- \in I^1(\omega)$ and $y^- \rightharpoonup y^+ \in \mathcal{R}$, then $y^+ \in I^1(\omega)$.
— for all $y \in I^1(\omega)$, if $y \neq y_*^-$, then $y_2 < y_{2,*}^-$.

Let $(Z(t))$ be a birth and death process on $\mathbb{N}$ starting at 0, with the transitions

$$x \to x+ \begin{cases} +k_m & \lambda, \\ -1 & \mu \text{ if } x \geq 1, \end{cases}$$

for

$$\lambda \overset{\text{def.}}{=} \sum_{\substack{r=y^-\rightharpoonup y^+\in\mathcal{R},\\ y^-\in I^1(\omega),\ y_2^-<y_2^+}} \kappa_r M_1^{y_1^-} M_2^{\omega_2(k_m-1)} \quad \text{and} \quad \mu \overset{\text{def.}}{=} \kappa_{r_*} m_1^{y_{*,1}^-} \left(\frac{M_2}{2}\right)^{\omega_2 k_m}.$$

The process $(Z(t))$ is almost an $M/M/1$ queue, with steps up equal to k_m, and steps down equal to -1. Choosing M_2 large enough, depending on m_1 and M_1, we can make the assumption that $k_m\lambda < \mu$.

Using a similar argument as the argument used in the proof of Proposition 6 in Laurence and Robert [53], we can construct a coupling of $(X_2^N(t))$ and $(Z(t))$ such that the relation

$$X_2^N(t \wedge \tau_N) \leq \left(\frac{M_2}{2}N\right)^{\omega_2} + Z((t \wedge \tau_N)N^{\rho_0(\omega)}), \quad \forall t \geq 0,$$

holds. This leads to

$$\mathbb{P}_{x_N}\left(\sup_{t\leq t_0} \overline{X}_2^N(s \wedge H_1^N) \ \geq M_2^{\omega_2}\right) \leq \mathbb{P}\left(\sup_{t\leq t_0} Z(tN^{\rho_0(\omega)-\gamma}) \ \geq \left(\frac{M_2}{2}N\right)^{\omega_2}\right),$$

and using Proposition 5.15 of Appendix 7, we can conclude for the upper-bound. The lower-bound is shown similarly.

Control of H_{1+}^N: Upper bound of $(\overline{X}_1^N(t))$. For all $t \geq 1$, $X_1^N(t) \leq S^N(t)$, and $(S^N(t))$ is solution of the SDE

$$\mathrm{d}S^N(t) = \sum_{r=y^-\rightharpoonup y^+\in\mathcal{R}} \left\langle (y^+ - y^-), \omega \right\rangle \mathcal{P}_r\left(\left(0, \kappa_r (X_N(t-))^{(y^-)}\right), \mathrm{d}t\right). \tag{5.27}$$

in which all the positive terms verify $\langle y^-, \omega\rangle < \rho_d(\omega)$, see Proposition 5.4. Therefore,

$$\begin{aligned} &\sup_{t\leq t_0}\left(\overline{S}^N(t \wedge \overline{\tau}_N)\right) - \overline{S}^N(0) \\ &\leq \frac{1}{N} \sum_{\substack{r=y^-\to y^+\in\mathcal{R},\\ \langle y^-,\omega\rangle<\rho_d(\omega)}} \int_0^{t_0} \mathcal{P}_r\left(\left(0, \kappa_r M_1^{y_1^-} M_2^{\omega_2 y_2^-} N^{\langle y^-,\omega\rangle}\right), \frac{\mathrm{d}t}{N^{\rho_d(\omega)-1}}\right) \overset{\text{def.}}{=} R_N(t_0), \end{aligned} \tag{5.28}$$

and when N goes to infinity, $(R_N(t_0))$ goes to 0 almost surely. Taking $M_1 > 1$, this leads to

$$\lim_{N\to+\infty} \mathbb{P}(H_{1+}^N < \min\{t_0, H_{1-}^N, H_2^N\}) = 0.$$

Decrease of $(\overline{S}^N(t))$. Let $r_d = y_d^- \rightharpoonup y_d^+ \in \mathcal{R}$ a reaction such that $\langle y_d^-, \omega\rangle = \rho_d(\omega)$ and $\langle y_d^+ - y_d^-, \omega\rangle = -\delta_d < 0$. If $t_0 = \overline{\tau}_N$, integrating Relation (5.27), one gets :

$$\overline{S}^N(t_0) \le \overline{S}^N(0) + R_N(t_0) - \frac{\delta_d}{N}\int_0^{t_0} \mathcal{P}_{r_d}\left(\left(0, \kappa_{r_d} m_1^{y_1^-} m_2^{\omega_2 y_2^-} N^{\rho_d(\omega)}\right), \frac{\mathrm{d}t}{N^{\rho_d(\omega)-1}}\right), \tag{5.29}$$

and the right part of the inequality converges almost surely when N goes to infinity to

$$b(t_0) \stackrel{\text{def.}}{=} 1 - \delta_d t_0 \kappa_{r_d} m_1^{y_1^-} m_2^{\omega_2 y_2^-}.$$

Setting t_0 small enough so that $1 > b(t_0) > m_1$, we get

$$\limsup_{N\to+\infty} \sup_{(N,x_2^N)\in\mathcal{H}_c(\omega)} \mathbb{E}\left(\overline{S}^N(\overline{\tau}_N)\right) \le M_1 \lim_{N\to\infty} \mathbb{P}(H_{1+}^N \wedge H_2^N = \overline{\tau}_N) + b(t_0)$$

and using the previous results, this leads to Relation (5.26).

Proof of Proposition 5.12. We choose $\tau = \tau_N$. The second limit is straightforward, since $\gamma > -1$, and $\tau_N \le t_0/N^\gamma$.

If for $N \ge 1$, (y_1^N, y_2^N) is such that $y_1^N/N \in [m_1, M_1]$ and $y_2^N/N^{\omega_2} \in [m_2, M_2]$, using a simple Taylor expansion, one gets

$$\frac{V(y_N)}{N\ln(N)} = \frac{y_1^N}{N} + O(1/\ln(N)) = \frac{\langle y_N, \omega\rangle}{N} + O(1/\ln(N)),$$

and therefore,

$$\limsup_{N\to+\infty} \sup_{(N,x_2^N)\in\mathcal{H}_c(\omega)} \frac{\mathbb{E}\left(V(X_N(\tau_N))\right)}{N\ln(N)} = \limsup_{N\to+\infty} \sup_{(N,x_2^N)\in\mathcal{H}_c(\omega)} \mathbb{E}\left(\overline{S}^N(\overline{\tau}_N)\right)$$

which concludes the proof. □

6.3. The case of $\omega_2 = 1$.

Proposition 5.13. *If $\omega = (1,1) \in \mathcal{W}(\mathcal{X})$ and $c > 0$, one can find some stopping time τ, and some $a > 0$ such that*

$$\limsup_{N\to+\infty} \sup_{e^{-c}\le x_2^N/N\le 1} \left(\frac{\mathbb{E}_{(N,x_2^N)}(V(X(\tau)))}{N\ln(N)} - \frac{V(N,x_2^N)}{N\ln(N)}\right) \le -a \tag{5.30}$$

$$\text{and} \quad \limsup_{N\to+\infty} \sup_{e^{-c}\le x_2^N/N\le 1} \frac{\mathbb{E}_{(N,x_2^N)}(\tau)}{N} = 0.$$

Proof. The proof is very similar to the proof of Proposition 5.12, except that $(X_2^N(t))$ does not have to be uniformly bounded as before. We keep the notations defined in the previous proof, with $\omega_2 = 1$. We have

$$S^N(t) = X_1^N(t) + X_2^N(t).$$

Relation 5.28 still holds until $\overline{\tau}_N$, and therefore, taking $M_1, M_2 > 2$, one gets easily

$$\lim_{N\to+\infty} \sup_{e^{-c}\le x_2^N/N\le 1} \mathbb{P}_{(N,x_2^N)}\left(H_{1+}^N \wedge H_{2+}^N = \overline{\tau}_N\right) = 0.$$

Similarly, Relation (5.29) still holds (with the same notations for reaction r_d and δ_d), and taking $m_1 < 1$, $m_2 < e^{-c}$ and $t_0 > 0$ such that

$$a(t_0) \stackrel{\text{def.}}{=} \delta_d t_0 \kappa_{r_d} m_1^{y_1^-} m_2^{\omega_2 y_2^-}$$

verify $a(t_0) < 1 - m_1$ and $a(t_0) < e^{-c} - m_2$, we get

$$\limsup_{N\to+\infty} \sup_{e^{-c}\le x_2^N/N\le 1} \left(\mathbb{E}_{(N,x_2^N)}(\overline{S}^N(\overline{\tau}_N)) - 1 - \frac{x_2^N}{N} \right) \le -a(t_0). \tag{5.31}$$

If for $N\geq 1$, (y_1^N, y_2^N) is such that $y_1^N/N \in [m_1, M_1]$ and $y_2^N/N \in [m_2, M_2]$, using a simple Taylor expansion, one gets

$$\frac{V(y_N)}{N\ln(N)} = \frac{y_1^N + y_2^N}{N} + O(1/\ln(N)) = \frac{\langle y_N, \omega\rangle}{N} + O(1/\ln(N)),$$

and therefore, Relation (5.31) leads to the first inequality of Proposition 5.13, the second inequality being straightforward. □

6.4. Practical corollary. The following corollary is a direct consequence of Propositions 5.12 and 5.13.

COROLLARY 5.14. *If $\omega \in \mathcal{W}(\mathcal{X})$, and $c > 0$, one can find some stopping time τ, such that for N large enough, for all $(N, x_2^N) \in \mathcal{H}_c(\omega)$ such that $x_2^N \le N$,*

$$\mathbb{E}_{(N,x_2^N)}(V(X(\tau))) - V(x_N) \le -\mathbb{E}_{(N,x_2^N)}(\tau). \tag{5.32}$$

7. Conclusion

PROOF OF THEOREM 5.1. Proposition 5.11 gives us some $K_1 \geq K_2 \geq 1$, and some $c_\omega > 0$ for each $\omega \in \mathcal{W}(\mathcal{X})$) such that Relation (5.4) holds for x in

$$B(K_1, K_2) \setminus \left(\bigcup_{\omega\in\mathcal{W}(\mathcal{X})} \mathcal{H}_{c_\omega}(\omega) \right),$$

and $\tau = t_1$ the first instant of jump of the process, defined in Relation (5.9).

For the states in $\mathcal{H}_{c_\omega}(\omega)$, for each of the $\omega \in \mathcal{W}(\mathcal{X})$, we use Corollary 5.14. K_2 is not changed, and we choose K_1 large enough so that if $N \geq K_1/2$, Relation (5.32) holds. This concludes the proof of Theorem 5.1. □

Appendix

A technical result: upper bound of an $M/M/1$ queue.

LEMMA 5.15. *Let $p \geq 1$. If $(Z(t))$ is a process starting at 0 and with transitions*

$$x \mapsto \begin{cases} x + p & \text{with rate} \quad \lambda \\ x - 1 & \text{with rate} \quad \mu, \end{cases} \tag{5.33}$$

with $0 < \lambda p < \mu$, for $\alpha, t_0 > 0$, the following limit holds:

$$\lim_{N\to+\infty} \mathbb{P}\left(\sup_{t<t_0} \{Z(tN^\alpha)\} \geq N \right) = 0. \tag{5.34}$$

PROOF. Let $\mathcal{N}_\lambda$ and $\mathcal{N}_\mu$ be independent Poisson processes with respective intensities $\lambda\,\mathrm{d}t$ on $\mathbb{R}_+$ and $\mu\,\mathrm{d}t$ on $\mathbb{R}_+$, and set for $t\geq 0$

$$S(t) = p\mathcal{N}_\lambda((0,t]), -\mathcal{N}_\mu((0,t]),$$

First, lets show that

$$\delta_c \stackrel{\text{def.}}{=} \mathbb{P}\left(\sup_{s\geq 0} S(s) \geq 1 \right) < 1.$$

The ergodic theorem implies that $(S(t)/t)$ is converging to $\lambda p - \mu < 0$, almost surely. If δ_c is 1, then the variable

$$T_1 = \inf\{t\geq 0 : S(t) \geq 1\}$$

is almost surely finite. Since $(S(t+T_1)-S(T_1))$ has the same distribution as $(S(t))$, one concludes that the lim sup of $(S(t))$ is $+\infty$. Contradiction.

Now for $N \geq 1$, let

$$H_N \stackrel{\text{def.}}{=} \inf\{t : Z(t) \geq N\}.$$

Lets show that we can find $\alpha \in (0,1)$ such that $(\alpha^N H_N)$ converges in distribution to $+\infty$. This is enough to conclude to the result of Lemma 5.15.

As long as $(X(t))$ does not hit 0, the process behaves as the random walk $(S(t))$. Hence, starting from $x \geq 0$, the process $(X(t))$ returns to 0 before going above $x+1$ with a probability at least $1-\delta_c > 0$. The duration of this step is stochastically lower bounded by an exponential random variable with parameter $\lambda+\mu$.

To reach N starting from 0, we must have at least $N_p = \lfloor N/p \rfloor$ consecutive positive jumps in a row without returning to 0. Let (B_i) a sequence of i.i.d. Bernoulli random variables with parameter δ_c and, for $n \geq 1$,

$$\nu_n = \inf\{k \geq n : B_{k-n+1} = B_{k-n+2} = \cdots = B_k = 1\}$$

For $k \in \mathbb{N}$, we have

$$\mathbb{P}(\nu_n \leq k) \leq k\delta_c^n.$$

We denote by (F_i) an i.i.d. sequence of exponential random variables with parameter $\lambda+\mu$. By using the fact that F_1 has a finite positive exponential moment, for $w_0 < 1/(\lambda+\mu)$, there exist constants, C_0, $\alpha_1 \in (0,1)$ and $M_0 \geq 0$, such that for $M \geq M_0$,

$$\mathbb{P}\left(\frac{1}{M}\sum_{i=1}^{M} F_i \leq w_0\right) \leq C_0 \alpha_1^M.$$

For $\alpha_2 \in (0,1)$,

$$\begin{aligned}\mathbb{P}\left(\alpha_2^N H_N \leq T\right) &\leq \mathbb{P}\left(\alpha_2^N \sum_{i=1}^{\nu_{N_p}} F_i \leq T\right) \\ &\leq \mathbb{P}\left(\frac{1}{K_0\lceil \alpha_2^{-N}\rceil} \sum_{i=1}^{K_0\lceil \alpha_2^{-N}\rceil} F_i \leq \frac{1}{K_0 \alpha_2^N \lceil \alpha_2^{-N}\rceil} T\right) + \mathbb{P}\left(\nu_{N_p} \leq K_0\lceil \alpha_2^{-N}\rceil\right) \\ &\leq C_0 \alpha_1^{K_0\lceil \alpha_2^{-N}\rceil} + K_0\lceil \alpha_2^{-N}\rceil \delta_c^{N_p}.\end{aligned}$$

If we fix $\alpha_2 \in (\delta_c^{1/p}, 1)$, and N_0 so large that if $N \geq N_0$, then $\alpha_2^N \lceil \alpha_2^{-N}\rceil \geq 1/2$ and $\alpha_2^{-N} > N \vee M_0$, and, finally $K_0 \geq 1$ large enough so that $2T/K_0 \leq w_0$, then we get a constant $\alpha_3 \in (0,1)$ and $C_1 \geq 0$, such that for $N \geq N_0$, the relation

$$\mathbb{P}\left(\alpha_2^N H_N \leq T\right) \leq C_1 \alpha_3^N$$

holds. The proposition is proved. □

Proofs from Section 3. We will here give a sketch of the proofs of the propositions in Section 3.

Proposition 5.4 is a consequence of the weak reversibility of the CRN. The proof would be tedious to write but presents no difficulty.

For Proposition 5.6, $\mathcal{X}$ has one single linkage class.

(a) Let $(y, y') \in A \times \mathcal{C} \setminus A$. Since the CRN is weakly reversible and has a single linkage class, there is a path from y to y', and therefore there is necessarily a reaction $y^- \rightharpoonup y^+ \in \mathcal{R}$ such that $y^- \in A$ and $y^+ \notin A$.
(b) If $I^d(\omega) \neq \emptyset$, we can't have $I^1(\omega) = \mathcal{C}$, and using (a), we can conclude.
(c) If $I^1(\omega) \cap I^d(\omega) = \emptyset$, then $I^d(\omega) = \emptyset$, and therefor $I^1(\omega) = \mathcal{C}$. Moreover, if $y \neq y'$, since $\langle \omega, y\rangle = \langle \omega, y'\rangle$, we have $y_1 - y_1' = \omega_2(y_2' - y_2)$, and therefore $y_2 \neq y_2'$.

(d) If $\omega, \omega' \in \mathcal{W}(\mathcal{X})$, because of (c), for $y, y' \in \mathcal{C}$, $\langle \omega, y' - y \rangle = 0 = \langle \omega', y' - y \rangle$. Note that we still have $\omega_1 = \omega'_1 = 1$. Therefore, $\omega_2(y'_2 - y_2) = \omega'_2(y'_2 - y_2)$, which leads to $\omega_2 = \omega'_2$.

For Proposition 5.8, lets do the proof by contradiction. If for some $\omega_2 \in (0, 1]$, setting $\omega = (1, \omega_2)$, $I^d(\omega) = \emptyset$, then for each reaction $y^- \rightharpoonup y^+ \in \mathcal{R}$, $\langle \omega, y^+ - y^- \rangle = 0$. Therefore, $(X(t))$ verify a mass conservation equation, and $(x + \mathbb{S}) \cap \mathbb{N}^2$ is finite.

Proposition 5.9 is straightforward.

Bibliography

[1] A. Agazzi, A. Dembo, and J.-P. Eckmann. "On the Geometry of Chemical Reaction Networks: Lyapunov Function and Large Deviations". In: *Journal of Statistical Physics* 172.2 (2018), pp. 321–352.

[2] A. Agazzi, D. F. Anderson, D. Cappelletti, and J. C. Mattingly. "Solving the chemical recurrence conjecture in 2D". In: *Stochastic Networks Conference*. Cornell University, June 2022.

[3] A. Agazzi, A. Dembo, and J.-P. Eckmann. "Large deviations theory for Markov jump models of chemical reaction networks". In: *The Annals of Applied Probability* 28.3 (2018), pp. 1821–1855.

[4] A. Agazzi and J. C. Mattingly. "Seemingly stable chemical kinetics can be stable, marginally stable, or unstable". In: *Communications in Mathematical Sciences* 18.6 (2020), pp. 1605–1642.

[5] D. F. Anderson. "A proof of the global attractor conjecture in the single linkage class case". In: *SIAM Journal on Applied Mathematics* 71.4 (2011), pp. 1487–1508.

[6] D. F. Anderson, G. Craciun, M. Gopalkrishnan, and C. Wiuf. "Lyapunov functions, stationary distributions, and non-equilibrium potential for reaction networks". In: *Bulletin of mathematical biology* 77.9 (2015), pp. 1744–1767.

[7] D. F. Anderson, G. Craciun, and T. G. Kurtz. "Product-form stationary distributions for deficiency zero chemical reaction networks". In: *Bulletin of mathematical biology* 72.8 (2010), pp. 1947–1970.

[8] D. F. Anderson and J. Kim. "Some network conditions for positive recurrence of stochastically modeled reaction networks". In: *SIAM Journal on Applied Mathematics* 78.5 (2018), pp. 2692–2713.

[9] D. F. Anderson, D. Cappelletti, and J. Kim. "Stochastically modeled weakly reversible reaction networks with a single linkage class". 2019.

[10] D. F. Anderson, D. Cappelletti, J. Kim, and T. D. Nguyen. "Tier structure of strongly endotactic reaction networks". In: *Stochastic Processes and their Applications* 130.12 (2020), pp. 7218–7259.

[11] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala. "Asymptotic analysis of multiscale approximations to reaction networks". In: *The Annals of Applied Probability* 16.4 (2006), pp. 1925–1961.

[12] J. Bertoin and M. Yor. "Exponential functionals of Lévy processes". In: *Probability Surveys* 2 (2005), pp. 191–212.

[13] E. Bibbona, J. Kim, and C. Wiuf. "Stationary distributions of systems with discreteness-induced transitions". In: *Journal of The Royal Society Interface* 17.168 (2020), p. 20200243.

[14] P. Billingsley. *Convergence of probability measures*. Second Edition. Wiley Series in Probability and Statistics: Probability and Statistics. New York: John Wiley & Sons Inc., 1999, pp. x+277.

[15] A. Bovier and F. den Hollander. *Metastability*. Vol. 351. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2015, pp. xxi+581.

[16] M. Bramson. “Instability of FIFO queueing networks with quick service times”. In: *Annals of Applied Probability* 4.3 (1994), pp. 693–718.

[17] M. Bramson. *Stability of queueing networks*. Vol. 1950. Lecture Notes in Mathematics. Springer, Berlin, 2008, pp. viii+190.

[18] D. Cappelletti and B. Joshi. “Graphically Balanced Equilibria and Stationary Measures of Reaction Networks”. In: *SIAM Journal on Applied Dynamical Systems* 17.3 (2018), pp. 2146–2175.

[19] D. Cappelletti and C. Wiuf. “Product-Form Poisson-Like Distributions and Complex Balanced Reaction Systems”. In: *SIAM Journal on Applied Mathematics* 76.1 (2016), pp. 411–432.

[20] R. Cowan and S. N. Chiu. “A Stochastic Model of Fragment Formation When DNA Replicates”. In: *Journal of Applied Probability* 31.2 (1994), pp. 301–308.

[21] D. A. Dawson. “Measure-valued Markov processes”. In: *École d’Été de Probabilités de Saint-Flour XXI—1991*. Vol. 1541. Lecture Notes in Math. Berlin: Springer, Nov. 1993, pp. 1–260.

[22] C. del Sole. “Stochastic models of Chemical Reaction Networks: multiscale approximations and convergence”. MA thesis. Politecnico di Torino, 2020.

[23] A. K. Erlang. “The Theory of Probabilities and Telephone Conversations”. In: *Nyt Tidsskrift for Matematik B* 20 (1909).

[24] S. Ethier and T. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 2009.

[25] M. Feinberg. *Lectures on chemical reaction networks, written version of lectures given at the Mathematical Research Center, University of Wisconsin, Madison, WI, 1979*. 1979.

[26] M. Feinberg. “Complex balancing in general kinetic systems”. In: *Archive for rational mechanics and analysis* 49.3 (1972), pp. 187–194.

[27] M. Feinberg. *Foundations of chemical reaction network theory*. Vol. 202. Applied Mathematical Sciences. Springer, Cham, 2019, pp. xxix+454.

[28] M. Feinberg. “The existence and uniqueness of steady states for a class of chemical reaction networks”. In: *Archive for Rational Mechanics and Analysis* 132.4 (1995), pp. 311–370.

[29] M. Feinberg and F. Horn. “Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspaces”. In: *Archive for Rational Mechanics and Analysis* 66 (1977), pp. 83–97.

[30] Y. Filonov. “A criterion for the ergodicity of discrete homogeneous Markov chains”. In: *Akademiya Nauk Ukrainskoi SSR. Institut Matematiki. Ukrainskii Matematicheskii Zhurnal* 41.10 (1989), pp. 1421–1422.

[31] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*. Second. New York: Springer-Verlag, 1998, pp. xii+430.

[32] V. Fromion, P. Robert, and J. Zaherddine. “A Stochastic Analysis of Networks with Pairing”. In: *Stochastic Processes and their Applications* 178 (Dec. 2024), p. 104480.

[33] C. Gallinger and L. Popovic. “Asymmetric autocatalytic reactions and their stationary distribution”. Nov. 2023.

[34] F. Guillemin, P. Robert, and B. Zwart. “AIMD algorithms and exponential functionals”. In: *The Annals of Applied Probability* 14.1 (2004), pp. 90–117.

[35] C. M. Guldberg and P. Waage. “Studies concerning affinity”. In: *CM Forhandlinger: Videnskabs-Selskabet i Christiana* 35.1864 (1864), p. 1864.

[36] J. Gunawardena. “Chemical reaction network theory for in-silico biologists”. In: *Notes available at http://vcp. med. harvard. edu/papers/crnt. pdf* (2003), p. 5.

[37] M. Hairer. *Convergence of Markov processes*. 2010.

[38] J. M. Hammersley. "On counters with random dead time. I". In: *Proceedings of the Cambridge Philosophical Society* 49 (1953), pp. 623–637.

[39] R. Z. Has'minskiĭ. "On the principle of averaging the Itô's stochastic differential equations". In: *Kybernetika (Prague)* 4 (1968), pp. 260–279.

[40] R. Z. Has'minskiĭ. "Principle of Averaging for Parabolic and Elliptic Differential Equations and for Markov Processes with Small Diffusion". In: *Theory of Probability & Its Applications* 8.1 (1963), pp. 1–21.

[41] L. Hoessly and C. Mazza. "Stationary Distributions and Condensation in Autocatalytic Reaction Networks". In: *SIAM Journal on Applied Mathematics* 79.4 (2019), pp. 1173–1196.

[42] F. Horn and R. Jackson. "General mass action kinetics". In: *Archive for rational mechanics and analysis* 47.2 (1972), pp. 81–116.

[43] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*. Second. Vol. 288. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 2003, pp. xx+661.

[44] A. Jakubowski. "A Non-Skorohod Topology on the Skorohod Space". In: *Electronic Journal of Probability* 2 (1997), pp. 1–21.

[45] C. Jia, D.-Q. Jiang, and Y. Li. "Detailed balance, local detailed balance, and global potential for stochastic chemical reaction networks". In: *Advances in Applied Probability* 53.3 (2021), pp. 886–922.

[46] H.-W. Kang and T. G. Kurtz. "Separation of time-scales and model reduction for stochastic reaction networks". In: *The Annals of Applied Probability* 23.2 (2013), pp. 529–583.

[47] H.-W. Kang, T. G. Kurtz, and L. Popovic. "Central limit theorems and diffusion approximations for multiscale Markov chain models". In: *The Annals of Applied Probability* 24.2 (2014), pp. 721–759.

[48] F. P. Kelly. *Reversibility and stochastic networks*. Chichester: John Wiley & Sons Ltd., 1979, pp. viii+230.

[49] J. Kern. "Skorokhod topologies: What they are and why we should care". In: *Mathematische Semesterberichte* 71.1 (Oct. 2023), pp. 1–18.

[50] J. K. Kim, G. A. Rempala, and H.-W. Kang. "Reduction for Stochastic Biochemical Reaction Networks with Multiscale Conservations". In: *Multiscale Modeling & Simulation* 15.4 (2017), pp. 1376–1403.

[51] T. G. Kurtz. "The Relationship between Stochastic and Deterministic Models for Chemical Reactions". In: *Journal of Chemical Physics* 57.7 (Oct. 1972), pp. 2976–2978.

[52] L. Laurence and P. Robert. "A Scaling Approch to Stochastic Chemical Reaction Networks". Sept. 2024.

[53] L. Laurence and P. Robert. "Analysis of Stochastic Chemical Reaction Networks with a Hierarchy of Timescales". Aug. 2024.

[54] L. Laurence and P. Robert. "Stochastic Chemical Reaction Networks with Discontinuous Limits and AIMD processes". June 2024.

[55] T. Kurtz. "Averaging for martingale problems and stochastic approximation". In: *Applied Stochastic Analysis, US-French Workshop*. Vol. 177. Lecture notes in Control and Information sciences. Springer Verlag, 1992, pp. 186–209.

[56] E. W. Lund. "Guldberg and Waage and the law of mass action". In: *Journal of Chemical Education* 42.10 (1965), p. 548.

[57] J. K. McSweeney and L. Popovic. "Stochastically-induced bistability in chemical reaction systems". In: *The Annals of Applied Probability* 24.3 (2014), pp. 1226–1268.

[58] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Second. Cambridge University Press, Cambridge, 2009, pp. xxviii+594.

[59] A. Mielke, R. I. A. Patterson, M. A. Peletier, and D. R. Michiel Renger. "Non-equilibrium Thermodynamical Principles for Chemical Reactions with Mass-Action Kinetics". In: *SIAM Journal on Applied Mathematics* 77.4 (2017), pp. 1562–1585.

[60] P. Mozgunov, M. Beccuti, A. Horvath, T. Jaki, R. Sirovich, and E. Bibbona. "A review of the deterministic and diffusion approximations for stochastic chemical reaction networks". In: *Reaction Kinetics, Mechanisms and Catalysis* 123.2 (2018), pp. 289–312.

[61] F. Kelly. "Loss Networks". In: *Annals of Applied Probability* 1.3 (1991), pp. 319–378.

[62] F. Kelly. "Blocking probabilities in large circuit-switched networks". In: *Advances in Applied Probability* 18 (1986), pp. 473–505.

[63] G. Papanicolaou, D. W. Stroock, and S. R. S. Varadhan. "Martingale approach to some limit theorems". In: *Proc. 1976. Duke Conf. On Turbulence.* III. Duke Univ. Math, 1977.

[64] J. F. C. Kingman. *Poisson processes.* Oxford studies in probability, 1993.

[65] L. Popovic. "Large deviations of Markov chains with multiple time-scales". In: *Stochastic Processes and their Applications* 129.9 (2019), pp. 3319–3359.

[66] H. Chen and A. Mandelbaum. "Discrete flow networks: bottleneck analysis and fluid approximations". In: *Mathematics of Operation Research* 16.2 (May 1991), pp. 408–446.

[67] P. Robert. *Stochastic Networks and Queues.* Vol. 52. Stochastic Modelling and Applied Probability Series. New-York: Springer, June 2003, pp. xix+398.

[68] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000, pp. xx+386.

[69] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000, pp. xiv+480.

[70] A. N. Rybko and A. L. Stolyar. "On the ergodicity of random processes that describe the functioning of open queueing networks". In: *Problems on Information Transmission* 28.3 (1992), pp. 3–26.

[71] N. Saito and K. Kaneko. "Theoretical analysis of discreteness-induced transition in autocatalytic reaction dynamics". In: *Phys. Rev. E* 91 (2 Feb. 2015), p. 022707.

[72] M. Sharpe. *General theory of Markov processes.* Vol. 133. Pure and Applied Mathematics. Academic Press, Inc., Boston, MA, 1988, pp. xii+419.

[73] Y. Togashi and K. Kaneko. "Switching dynamics in reaction networks induced by molecular discreteness". In: *Journal of Physics: Condensed Matter* 19.6 (Jan. 2007), p. 065150.

[74] Y. Togashi and K. Kaneko. "Transitions Induced by the Discreteness of Molecules in a Small Autocatalytic System". In: *Phys. Rev. Lett.* 86 (11 Mar. 2001), pp. 2459–2462.

[75] F. Verhulst. *Nonlinear differential equations and dynamical systems.* Second. Universitext. Springer-Verlag, Berlin, 1996, pp. x+303.

[76] E. O. Voit, H. A. Martens, and S. W. Omholt. "150 years of the mass action law". In: *PLoS Comput Biol* 11.1 (2015), e1004012.

[77] W. Whitt. *Stochastic-process limits.* Springer Series in Operations Research. Springer-Verlag, New York, 2002, pp. xxiv+602.

[78] E. T. Whittaker and G. N. Watson. *A course of modern analysis.* Cambridge Mathematical Library. Cambridge: Cambridge University Press, 1996, pp. vi+608.

[79] C. Xu, M. C. Hansen, and C. Wiuf. "Full classification of dynamics for one-dimensional continuous time Markov chains with polynomial transition rates". In: *Advances in Applied Probability* 55.1 (2023), pp. 321–355.

[80] M. Yor. *Exponential functionals of Brownian motion and related processes.* Springer Finance. Springer-Verlag, Berlin, 2001, pp. x+205.

Remerciements

Merci Philippe pour les discussions culinaires, de littérature, bandes dessinées, vin, politique et j'en passe. Merci également, et évidemment, pour ton encadrement pendant ces quelques années, pour ta patience, ton soutien sans faille, ta disponibilité, et pour m'avoir introduit aux réseaux chimiques.

Merci Jana pour les fous rires, les gâteaux libanais, et le soutien moral. Merci Gaëtan pour les conseils, sur la thèse et de lecture.

Merci à Daniele Cappelletti et ses collègues italiens de nous avoir chaleureusement accueillis dans la communauté des chercheurs des réseaux chimiques.

Je suis très reconnaissante à Léa Popovic et Carl Graham d'avoir accepté d'être les rapporteurs de ma thèse. Je remercie également Amandine Véber, Marie Doumic, Bénédicte Haas et Irina Kourkova de prendre le temps de participer à mon jury de thèse.

Enfin, merci Chloé, Mathilde, Marie, Marion, et tous·tes les autres, pour les nécessaires moments de détente.