



Center for Research and Advanced  
Studies of the National  
Polytechnic Institute  
Zacatenco Research Center

Department of Mathematics

## **Stochastic Modeling of Chemical Reactions with Delay**

A thesis presented by

**José Eduardo Márquez Prado**

to obtain the degree of

**Master of Science in Mathematics**

Advisor:

**Dr. Eduardo Santillan Zeron**



Centro de Investigación y de  
Estudios Avanzados del Instituto  
Politécnico Nacional  
Unidad Zacatenco

Departamento de Matemáticas

## **Modelado Estocástico de Reacciones Químicas con Retardo**

Tesis presentada por

**José Eduardo Márquez Prado**

para obtener el grado de

**Maestro en Ciencias en la especialidad de Matemáticas**

Asesor de Tesis:

**Dr. Eduardo Santillan Zeron**

## ABSTRACT

In this thesis, we analyze the dynamics of a system of chemical reactions as a continuous-time stochastic process. In particular, when in the system only instantaneous chemical reactions occur we will see that the corresponding stochastic process is a continuous-time Markov chain, which is the unique solution of the Anderson-Kurtz's stochastic equation. We present an elementary proof of this last fact since we show that the Doob-Gillespie's algorithm and Anderson-Kurtz's algorithm are equivalent.

Finally, we will see how a simple modification to Anderson-Kurtz's stochastic equation allows us to model chemical systems where one of the reactions has a time delay. We will compare the corresponding Anderson-Kurtz's algorithm for this kind of chemical system with Barrio's algorithm, which is one of the most used tools to simulate chemical systems with delays.

## RESUMEN

En esta tesis analizamos la dinámica de un sistema de reacciones químicas como un proceso estocástico a tiempo continuo. En particular, cuando en el sistema solo ocurren reacciones químicas instantáneas, veremos que el correspondiente proceso estocástico es una cadena de Markov a tiempo continuo, la cual es la única solución de la ecuación estocástica de Anderson-Kurtz. Presentamos una prueba elemental de este último hecho, ya que demostramos que el algoritmo de Doob-Gillespie y el algoritmo de Anderson-Kurtz son equivalentes.

Finalmente, veremos como una simple modificación a la ecuación estocástica de Anderson-Kurtz nos permite modelar sistemas químicos donde una de las reacciones tiene retardo. Compararemos el correspondiente algoritmo de Anderson-Kurtz para esta clase de sistema químicos con el algoritmo de Barrio, el cual es una de las herramientas más usadas para simular sistemas químicos con retardos.

## AGRADECIMIENTOS

Esta tesis está dedicada a las siguientes personas:

Raúl Márquez Aguilar

Antonia Elia Prado Espinoza

Isabel Aguilar Ramos

David Prado Orozco

También, agradezco profundamente al Departamento de Matemáticas del Cinvestav por todo el apoyo ofrecido durante mi maestría.

Finalmente, agradezco a Conacyt por financiar este tipo de proyectos.

## CONTENTS

1. Introduction	2
2. Chemical Systems and Probability	6
2.1. Basic definitions and Fundamental Hypothesis	6
2.2. Construction of the reaction jump process	8
2.3. Anderson-Kurtz's Equation	11
2.4. Backward and Forward Kolmogorov's Equations	14
2.5. Chemical Master Equation(CME)	19
3. Chemicals Reactions with Delays	23
3.1. Reaction Jump Process of a Chemical System with one Delayed Reaction	24
3.2. A Semi-Markov Chemical Reaction System	27
4. Simulation of the Reaction Jump Process	32
4.1. Doob-Gillespie Stochastic Simulation Algorithm (SSA)	32
4.2. Anderson-Kurtz's SSA	38
5. Simulation of the Reaction Jump Process with Delays	48
5.1. Anderson-Kurtz's Stochastic Simulation Algorithm with Delays (SSAD)	48
5.2. Barrio's SSAD	53
5.3. Example	57
5.4. Cai's SSAD	63
6. Conclusions	69
References	71

## 1. INTRODUCTION

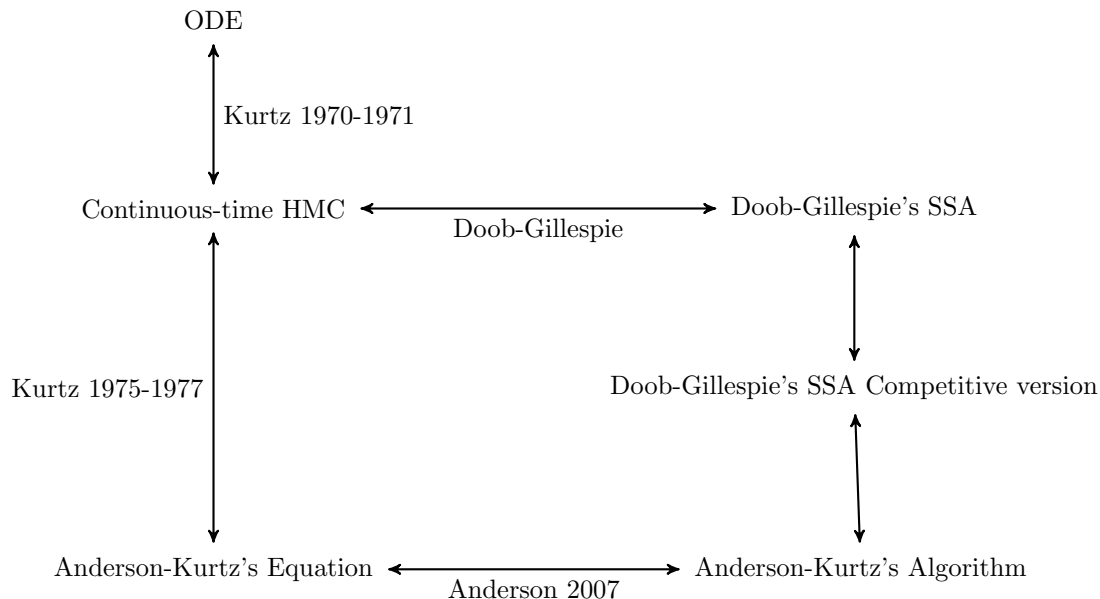
The main objective of this thesis work is to analyze a system of chemical reactions as a stochastic process, especially in the case when one of the chemical reactions has a time delay. A chemical reaction system or network comprises a set of chemical species, a series of chemical reactions, and a collection of propensity functions. The set of species consists of the reactants and products of the chemical reactions, and the propensity functions have as domain a subset  $\mathbb{X} \subset \mathbb{N}_0^m$ , where  $m$  is the number of chemical species and  $\mathbb{X}$  is itself the domain of possible states of the chemical network. The simpler stochastic model for such networks treats the system as a continuous-time Homogeneous Markov Chain (HMC) with space of states  $\mathbb{X}$ , an infinitesimal generator determined by the propensity functions, and reactions modeled as possible transitions of the chain. Therefore, one of the techniques to study the dynamics of the chemical system is through algorithms that simulate the corresponding continuous-time HMC or reaction jump process. Among these algorithms, Doob-Gillespie's Stochastic Simulation Algorithm (SSA) is the fundamental tool used for this job; see for example [12].

Although Doob-Gillespie's SSA gives us a way to simulate the trajectories of the reaction jump process, it does not give us an explicit equation to represent the time dynamic of these trajectories. The work done by Kurtz to approximate a certain type of ODE by means of a sequence of continuous-time HMC (see [13], [14]) led him to show that the reaction jump process of a chemical system with  $n$  chemical reactions is the unique solution of the following stochastic integral equation (see [15], [16], [17]):

$$(1.1) \quad X(t) = x_0 + \sum_{j=1}^n \mathcal{U}_j \left( \int_0^t \alpha_j(X(s)) ds \right) v[j],$$

where  $\{\mathcal{U}_j\}_j$  is a sequence of independent and unit-rate Poisson processes, and  $\{\alpha_j\}_j$ ,  $\{v[j]\}_j$  are the sequences of propensity functions and stoichiometric vectors, respectively associated to the chemical reactions (in Chapter 2 we will present a formal definition of these sequences).

In [1], Anderson introduces the first algorithm to solve equation (1.1), which we call the Anderson-Kurtz's algorithm or Anderson-Kurtz's SSA. Due to the work done by Kurtz, it is clear that Anderson-Kurtz's algorithm and Doob-Gillespie's SSA are equivalent. One of our intentions in this thesis is to present a direct and elementary proof of this equivalence. To do this, we will introduce a modified version of Doob-Gillespie's SSA, which we will call Doob-Gillespie's SSA Competitive version (this name will be justified in Chapters 2 and 4). Then, we will show that this version of Doob-Gillespie's SSA is equivalent to Anderson-Kurtz's algorithm. The following diagram shows the different equivalences mentioned above.



On the other hand, various efforts have been made to model chemical systems with delayed reactions. Among these efforts, we can find algorithms that try to simulate the dynamics of these chemical systems; and two of these algorithms are widely used in the area of biochemistry, they are Barrio's Stochastic Simulation Algorithm with Delays (SSAD) and Cai's SSAD, which are modifications to Doob-Gillespie SSA; see for example [3] and [5]. The main problem with these algorithms is that, unlike Doob-Gillespie SSA, it is not clear whether they are implementations of some stochastic process that models the chemical system with delays. An effort in the direction of finding a stochastic process associated with a delayed system is made by Anderson in [1], where he shows how a simple modification to the Anderson-Kurtz's equation (1.1) can indeed model chemical systems with delays. For example, in Chapter 3 we will see that for a chemical system of  $n$  instantaneous chemical reactions and a delayed one, its corresponding modification to the Anderson-Kurtz's equation is given by

$$(1.2) \quad Y(t) = Y(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(Y(s)) ds \right) v[k] \\ + \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \omega_1 + \mathcal{U}_{de} \left( \int_0^{t-d} \alpha_{de}(Y(s)) ds \right) \omega_2.$$

Thus, a natural question arises: For a chemical system of  $n$  instantaneous chemical reactions and a delayed one, are the trajectories of the solution to equation (1.2) simulated through Barrio's SSAD and Cai's SSAD? Although we do not have an answer to this question, it is our belief that the answer is affirmative. In Chapter 5 we will introduce some results in this direction.



We must now introduce the basic definitions and assumptions that will allow us to model the dynamics of a system of  $n$  ( $n \in \mathbb{N}_0$ ) chemical reactions. Among these definitions and assumptions, the *Fundamental Hypothesis of the Stochastic Reaction Kinetics* and the *Well-Mixed Assumption* imply that we can model the dynamics of this chemical system with a continuous-time Homogeneous Markov Chain (HMC), which has a specific infinitesimal generator. From now on, we will refer to this continuous-time HMC as the reaction jump process. In Section 2.2 we will present a construction of this reaction jump process.

Thus, the problem of determining the dynamics of a system with  $n$  chemical reactions is reduced to a question of determining the transition probabilities or semigroup of the corresponding reaction jump process. In Section 2.4 and Section 2.5 we will verify that the semigroup of the reaction jump process satisfies Kolmogorov's equations. In particular, in Section 2.5 we will see that this semigroup satisfies the corresponding Forward Kolmogorov equation, also known as the Chemical Master Equation (CME). Therefore, the solutions of the CME give us the transition probabilities of the reaction jump process.

In Section 2.3 we will present Anderson-Kurtz's ideas for modeling a system of  $n$  chemical reactions. We will see that these ideas are based on the following fact: in order to determine the state of the system at a certain time  $t$ , we have to know how many times each reaction has happened until time  $t$ . We will see how the counting process of each chemical reaction is given, and how the dynamics of the chemical system are represented by the Anderson-Kurtz's equation (1.1).

Following the ideas in Section 2.1, we will begin Chapter 3 by presenting a definition of what a delayed reaction is; there we will see that a delayed reaction is constituted by a normal chemical reaction, which changes the stoichiometry of the system instantaneously and generates an excited complex  $Z$ . This latter complex will decompose itself and change the stoichiometry of the system  $d$  units of time later. Furthermore, in Section 3.1 we will see how equation (1.2) allows us to model the dynamics of a system with  $n$  chemical reactions and one delayed reaction. In section 3.2 we will present a simple example of a chemical system with only one delayed reaction; we will show that the corresponding reaction jump process is no longer a continuous-time HMC, it is instead a Semi-Markovian process.

Although we can study the dynamics of a chemical system (without delayed reactions) through the CME's solutions, in practice this can be difficult because the dimension of the system of differential equations can be very large (the dimension is equal to the cardinality of the set  $\mathbb{X}$  of possible states). For this reason, the dynamics of the chemical system is studied with algorithms that can simulate the corresponding reaction jump process. In Chapter 4 we will introduce

Doob-Gillespie's Stochastic Simulation Algorithm (SSA); we will see that Doob-Gillespie's SSA is just an implementation based on the fact that the reaction jump process is a continuous-time HMC. Following the construction of the reaction jump process given in Section 2.2, in Section 4.1 we will give an equivalent version of Doob-Gillespie's SSA, which has an interpretation as a competition between the different chemical reactions. In Section 4.2 we will introduce Anderson-Kurtz's SSA, which gives us a way to solve Anderson-Kurtz's equation; we will show that Anderson-Kurtz's SSA and the competitive version of Doob-Gillespie's SSA (given in Section 4.1) are equivalents. This immediately implies that the only solution to Anderson-Kurtz's equation is the corresponding reaction jump process simulated by the Doob-Gillespie's SSA.

Finally, in Chapter 5 we will present the most used algorithms that allow us to simulate the dynamics of a chemical system with  $n$  chemical reactions and one delayed reaction. We will begin this chapter by introducing Anderson-Kurtz's SSAD, which gives us a way to solve equation (1.2). In Section 5.7 we will present Barrio's SSAD for this chemical system with delays, and in Section 5.3 we will introduce an example of a chemical system that will allow us to demonstrate that although Anderson-Kurtz's SSAD and Barrio's SSAD have some differences, for this example the trajectories that Barrio's SSAD generates until certain time are solutions of the corresponding Anderson-Kurtz's equation with delays.

We will finish chapter 5 by presenting Cai's SSAD; we will see that unlike Barrio's SSAD and Anderson-Kurtz's SSAD that simulate all the times where a stoichiometric change occurs (either by a chemical reaction or the decomposition of an excited  $Z$  complex), Cai's SSAD only simulates the times where a chemical reaction occurs. We will see how the Cai's SSAD can be interpreted as a race between the chemical reactions in the system.

## 2. CHEMICAL SYSTEMS AND PROBABILITY

**2.1. Basic definitions and Fundamental Hypothesis.** Consider a system of molecules that belongs to different chemical species. These molecules move in a given environment, the (spatial) domain  $\mathbb{D}$ , where  $\mathbb{D} \subset \mathbb{R}^k$  for  $k \in \{1, 2, 3\}$ , and undergo chemical reactions which affect their population. We understand that a chemical reaction is a process where two or more molecules collide together, in the space  $\mathbb{D}$ , to form new chemical species; for example

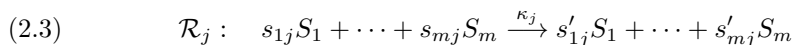


is a chemical reaction where one molecule of gaseous hydrogen reacts with a molecule of gaseous fluorine to produce two molecules of hydrogen fluoride.

As for the relation between spatial movement and reactivity, we make the following central assumption.

**Assumption 2.2.** (*Well-Mixed Assumption*) *The spatial movement of molecules is fast compared to reactions, i.e., the majority of close encounters of molecules (where molecules come close enough so that reactions are possible) are nonreactive and molecular positions are uniformly distributed throughout the space  $\mathbb{D}$  at any time.*

Given the well-mixed assumption, the spatial positions of molecules become insignificant for the system's dynamics, so that modeling approaches without spatial resolution are justified. The system's state may then be specified by counting the number of molecules of each species, regardless of their positions in the space. Hence, if in our chemical system there are  $m \in \mathbb{N}$  chemical species  $S_1, \dots, S_m$ , we only need to know the number of molecules of each species  $S_k$ . We do so by introducing a vector  $x \in \mathbb{N}_0^m$  so that the entry  $x_k$  is the number of molecules of the species  $S_k$ , for  $k = 1, 2, \dots, m$ . This vector  $x \in \mathbb{N}_0^m$  is called the state of the chemical system, and the set of all possible states  $x \in \mathbb{N}_0^m$  is denoted by  $\mathbb{X} \subset \mathbb{N}_0^m$ . A chemical reaction system or network comprises a set of chemical species, a set of chemical reactions, and a set of propensity functions. The set of species consists of the reactants and products of the chemical reaction, and the propensity functions have as domain a subset  $\mathbb{X} \subset \mathbb{N}_0^m$ , where  $m$  is the number of chemical species and  $\mathbb{X}$  is the set of possible states of the chemical network; for example, consider a system of  $m \in \mathbb{N}$  chemical species  $S_1, \dots, S_m$  which interact through  $n \in \mathbb{N}$  reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n$ . Each reaction is represented by a stoichiometric equation of the form



with the stoichiometric coefficients  $s_{lj}, s'_{lj} \in \mathbb{N}_0$  denoting the numbers of reactants and products molecules, respectively (see for example (2.1)). The associated stoichiometric vector

$$v[j] := (v_{1j}, \dots, v_{mj})^T \in \mathbb{Z}^m,$$

also called state-change vector, is defined as

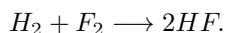
$$v_{lj} := s'_{lj} - s_{lj}$$

and describes the net change in the number of molecules of each species  $S_l$  due to reaction  $\mathcal{R}_j$ . The constant  $\kappa_j > 0$  is the reaction rate constant which quantifies the rate for the reaction to take place.

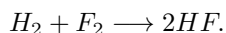
Each reaction  $\mathcal{R}_j$  acts on those  $x \in \mathbb{X}$  such that  $x + v[j] \in \mathbb{X}$ , in the following way

$$(2.4) \quad \mathcal{R}_j : x \rightarrow x + v[j].$$

For example, consider again the chemical reaction



Assume that the state of our system is given by the vector  $x \in \mathbb{N}_0^3$ , where  $x = (x_1, x_2, x_3)^T$  and  $x_1, x_2$ , and  $x_3$  represent the number of molecules of  $H_2, F_2$ , and  $HF$ , respectively; then the reaction



can be represented as

$$\mathcal{R} : x \rightarrow x + (-1, -1, 2)^T.$$

From now on we will only use the notation (2.4) to designate chemical reactions.

The probability for such a reaction to occur depends on the reaction rate constant and the state of the system.

**Definition 2.5.** (*Propensity function*) *The propensity function for the reaction  $\mathcal{R}_j$  is the function*

$$\alpha_j : \mathbb{X} \rightarrow [0, \infty)$$

that satisfies

$$(2.6) \quad \alpha_j(x) = \kappa_j h_j(x) \quad \forall x \in \mathbb{X},$$

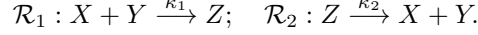
where  $h_j(x)$  is the number of distinct combinations in which the reactant molecules can participate into the reaction  $\mathcal{R}_j$ . By (2.3)

$$h_j(x) = \prod_{l=1}^m \binom{x_l}{s_{lj}}$$

provided that  $s_{lj} \leq x_l$  for all  $l = 1, \dots, m$ ; and  $h_j(x) = 0$  if for some  $1 \leq l \leq m$   $x_l < s_{lj}$ .

**Assumption 2.7.** (*Fundamental Hypothesis of the Stochastic Reaction Kinetics*)  
The probability for the reaction  $\mathcal{R}_j$  to occur once within the infinitesimal time interval  $[t, t + dt)$  is  $\alpha_j(x)dt$ , given that the system is in the state  $x$  at time  $t$ .

**Example 2.8.** (*Binding and Unbinding*) Consider the reaction system of three species  $X, Y$  and  $Z$  with reactions



For  $\mathcal{R}_1$  the stoichiometric vector and the propensity function are given by

$$v[1] = (-1, -1, 1)^T \quad \text{and} \quad \alpha_1(x, y, z) = \kappa_1 xy,$$

respectively; and for  $\mathcal{R}_2$

$$v[2] = (1, 1, -1)^T \quad \text{and} \quad \alpha_2(x, y, z) = \kappa_2 z.$$

**2.2. Construction of the reaction jump process.** Consider a reaction system of  $n$  reactions and  $m$  chemical species. To describe the time evolution of the system we consider the continuous-time stochastic process  $X = (X(t))_{t \geq 0}$ , where

$$X(t) = (X_1(t), \dots, X_m(t))^T \in \mathbb{X}$$

with  $X_j(t)$  denoting the number of molecules of the species  $S_j$  at time  $t \geq 0$  for  $j = 1, \dots, m$ . With this notation in mind, we can rewrite the fundamental hypothesis as

$$(2.9) \quad \mathbb{P}(X(t + dt) = x + v[j] | X(t) = x) = \alpha_j(x)dt$$

and the probability that a reaction doesn't occur in the time interval  $(t, t + dt]$  is given by

$$(2.10) \quad \mathbb{P}(X(t + dt) = x | X(t) = x) = 1 - \left( \sum_{j=1}^n \alpha_j(x) \right) dt$$

We are going to show the existence of a jump Markov chain that satisfies the equations above.

Let  $A$  be a real-valued square matrix of size  $|\mathbb{X}|^2$  ( $|\mathbb{X}| < \infty$ ) defined as follows

$$(2.11) \quad A(x, x + v[j]) = \alpha_j(x), \quad \forall \quad j = 1, 2, \dots, n,$$

$$(2.12) \quad A(x, x) = - \sum_{j=1}^n \alpha_j(x)$$

and

$$(2.13) \quad A(x, y) = 0 \quad \forall y \notin \{x + v[j] : 1 \leq j \leq n\}.$$

Observe that  $A$  satisfies

$$A(x, x) = - \sum_{y \in \mathbb{X} \setminus \{x\}} A(x, y)$$

and  $A(x, x) < \infty$ , that is  $A$  is a conservative and stable generator. We proceed with the construction of a jump Markov process with infinitesimal generator  $A$ .

Let  $X_0$  be a random variable with values in  $\mathbb{X}$ . If at time  $t = 0$  the chemical system is in the state  $X_0$ , then consider the following independent Homogeneous Poisson Processes (HPP)

$$\{N_{X_0, X_0+v[j]} : 1 \leq j \leq n\},$$

where each  $N_{X_0, X_0+v[j]}$  has intensity  $\alpha_j(X_0)$ . Let  $N_{X_0} = \sum_{j=1}^n N_{X_0, X_0+v[j]}$ . Then  $N_{X_0}$  is an HPP with intensity  $-A(X_0, X_0)$ , see for example [4]. If  $\tau_1$  is the time until the first event of  $N_{X_0}$  and  $J_1$  is the type of the event, then in the context of chemical reactions  $\tau_1$  is the time until the first reaction and  $J_1$  is the type of reaction. From the theory of HPP,  $\tau_1$  and  $J_1$  are independent random variables and

$$\mathbb{P}[\tau_1 > a, J_1 = j | X_0] = \exp(A(X_0, X_0)a) \frac{\alpha_j(X_0)}{-A(X_0, X_0)}$$

i.e.,  $\tau_1 \sim \exp(-A(X_0, X_0))$  and  $\tau_1 < \infty$  a.s. because  $A(X_0, X_0) < 0$  a.s.

Repeat the same process as above with  $X_1 = X_0 + v[J_1]$  and  $N_{X_1} = \sum_{j=1}^n N_{X_1, X_1+v[j]}$ , where given the event  $\{X(0) = x_0, X_1 = x_1\}$  the sets of HPP

$$\{N_{X_0, X_0+v[j]} : 1 \leq j \leq n\} \quad \{N_{X_1, X_1+v[j]} : 1 \leq j \leq n\}$$

are independent. Take random variables  $\tau_2, J_2$  such that

$$P[\tau_2 > a, J_2 = j | X_0, X_1] = \exp(A(X_1, X_1)a) \frac{\alpha_j(X_1)}{-A(X_1, X_1)}.$$

and let  $X_2 = X_1 + v[J_2]$ ,  $N_{X_2} = \sum_{j=1}^n N_{X_2, X_2+v[j]}$ .

By the process above, we get a sequence of r.v.  $\{X_i\}$  and  $\{t_i\}$  given by

$$X_i = X_0 + \sum_{k=1}^i v[J_k]$$

and

$$t_i = \sum_{k=1}^i \tau_k.$$

Observe that  $t_i$  is the time until the  $i$ -th reaction and the system will be in  $X_i$  at time  $t_i$ . Moreover,  $\forall x \in X$ ,  $A(x, x) < 0$  which implies that  $t_i < \infty$  a.s.

Now, by the construction of the sequences  $\{\tau_k\}$  and  $\{J_k\}$ , we can see

$$\mathbb{P}[\tau_{i+1} > a, J_{i+1}=j | X_0, \dots, X_i=x] = \mathbb{P}[\tau_{i+1} > a, J_{i+1}=j | X_i=x],$$

therefore

$$\begin{aligned} & \mathbb{P}[X_{i+1}=x+v[j], t_{i+1}-t_i > a | X_0, \dots, X_i=x] = \mathbb{P}[\tau_{i+1} > a, J_{i+1}=j | X_0, \dots, X_i=x] \\ & = \mathbb{P}[\tau_{i+1} > a, J_{i+1}=j | X_i=x] \\ & = \exp(A(x, x)a) \frac{\alpha_j(x)}{-A(x, x)}. \end{aligned}$$

By the continuity of the measure

$$\begin{aligned} & \mathbb{P}[X_{i+1}=x+v[j] | X_0, \dots, X_i=x] \\ & = \lim_{a \downarrow 0} P[X_{i+1}=x+v[j], t_{i+1}-t_i > a | X_0, \dots, X_i=x] \\ & = \lim_{a \downarrow 0} \exp(A(x, x)a) \frac{\alpha_j(x)}{-A(x, x)} \\ & = \frac{\alpha_j(x)}{-A(x, x)}. \end{aligned}$$

The argument above shows that  $\{X_i\}$  is a HMC with transitions probabilities

$$\mathbb{P}(X_{i+1}=x+v[j] | X_i=x) = \frac{\alpha_j(x)}{\sum_{j=1}^n \alpha_j(x)}$$

and

$$\mathbb{P}(X_{i+1}=y | X_i=x) = 0 \quad \forall y \notin \{x+v[j] : 1 \leq j \leq n\}.$$

With the discrete time HMC  $\{X_l\}_{l \in \mathbb{N}}$ , we can construct a continuous time stochastic process  $\{X(t)\}$  given by

$$X(t) = X_n \quad \text{if} \quad t_n \leq t < t_{n+1}$$

where  $t_0 = 0$  a.s. We say that  $\{X(t)\}$  is the reaction jump process of the chemical system and  $\{X_i\}_{i \in \mathbb{N}}$  is the corresponding embedded HMC. Observe that  $\{X_i\}_{i \in \mathbb{N}}$  is a HMC thanks to the hypothesis that  $A$  is conservative and stable. In general, as Doob showed (see [7]), the infinitesimal generator of a continuous-time HMC is stable and conservative if and only if with probability 1 the first discontinuity of the continuous-time HMC, after a specified time  $t$ , is a jump; that is, the infinitesimal generator of a continuous-time HMC is stable and conservative if and only if the embedded chain is a discrete HMC.

Since  $A$  is the infinitesimal generator of the reaction jump process  $\{X(t)\}$ , then when  $h \downarrow 0$

$$\begin{aligned} \mathbb{P}(X(t+h) = x + \nu[j] | X(t) = x) &= \alpha_j(x)h + o(h) \\ \mathbb{P}(X(t+h) = x | X(t) = x) &= 1 - \sum_{j=1}^n \alpha_j(x)h + o(h) \\ \mathbb{P}(X(t+h) = y | X(t) = x) &= 0 \quad \forall y \notin \{x + \nu[j] : 1 \leq j \leq n\} \end{aligned}$$

That is, the reaction jump process satisfies the Fundamental Hypothesis.

Consider the sequence of reactions times  $\{t_i\}$ . We define the explosion time of the chemical system as  $t_\infty := \lim_{i \rightarrow \infty} t_i$ . Note that given the event  $\{t_\infty < \infty\}$ , with probability 1 we have an infinite number of reactions in a small neighborhood of  $t_\infty$ , that is, the system jumps infinitely many times in a small neighborhood of  $t_\infty$ . Therefore, an important question about  $t_\infty$  is the following: under what conditions the event  $\{t_\infty = \infty\}$  has probability 1? Since the sequence  $\{t_i\}$  depends on the

generator  $A$ , then information about the generator  $A$  must give us information of  $t_\infty$ . But not only the properties of the generator are important because as Doob showed ([7]), there exist conservative and stable generators such that the jump process that defines it has finite explosion time with probability greater than 0. We're going to show that  $\{X(t)\}$  is regular, that is  $\mathbb{P}(t_\infty = \infty) = 1$ . First of all, we need Reuter's criterion (see [4]):

**Theorem 2.14.** (*Reuter's criterion*) *Let  $Q$  be a stable and conservative infinitesimal generator. Then the following are equivalent:*

(1) *The continuous-time HMC that  $Q$  defines is regular.*

(2) *For all  $\lambda > 0$  the system of equations*

$$(2.15) \quad (\lambda + q_i)y_i = \sum_{j \in E \setminus \{i\}} q_{i,j}y_j \quad i \in E,$$

*in matrix notation*

$$(2.16) \quad Qy = \lambda y,$$

*has the trivial solution  $y = 0$  as the only non-negative bounded solution.*

(3) *There is  $\lambda > 0$  such that the trivial solution  $y = 0$  is the only non-negative bounded solution of (2.15).*

Since  $A$  is a finite matrix, then the set of eigenvalues of  $A$  is finite, and there exist  $\lambda > 0$  that satisfies (3) of Theorem (2.14). Therefore  $\{X(t)\}$  is regular. We have assumed that  $\mathbb{X}$  is finite, but we can make a similar analysis when  $\mathbb{X}$  is a countable infinite set; we will show that the semigroup,  $\{P(t)\}$ , of the reaction jump process, satisfies the Forward Kolmogorov's equations (in our context the Chemical Master Equations)

$$P^{(1)}(t) = P(t)A,$$

which implies that at any time  $t_0 > 0$ , with probability 1 the last discontinuity of the reaction jump process before  $t_0$  is a jump (see [7]). Together with the fact that  $A$  is stable and conservative, the reaction jump process is regular, whether  $\mathbb{X}$  is finite or countable infinite. Therefore the name 'jump process' is justified and we can rewrite this process as

$$(2.17) \quad X(t) = \sum_{i=0}^{\infty} X_i \chi_{[t_i, t_{i+1})}(t).$$

**2.3. Anderson-Kurtz's Equation.** Another way to describe the jump process  $\{X(t)\}$  is with the help of a counting process  $R_j(t)$ , where  $R_j(t)$  is the number of reactions of type  $\mathcal{R}_j$  that have occurred in the time interval  $[0, t]$  for  $1 \leq j \leq n$ . Suppose that the chemical system satisfies  $X(0) = x_0$ . Then we can write the jump



process  $X(t)$  as:

$$(2.18) \quad X(t) = x_0 + \sum_{j=1}^n R_j(t)v[j],$$

and the fundamental hypothesis as:

**Assumption 2.19.** (*Fundamental Hypothesis of the Stochastic Reaction Kinetics*).  
For all  $j = 1, \dots, m$  and  $t \geq 0$  it holds

$$\begin{aligned} \mathbb{P}(R_j(t+dt) - R_j(t) = 0 | X(t) = x) &= 1 - \alpha_j(x)dt + o(dt), \\ \mathbb{P}(R_j(t+dt) - R_j(t) = 1 | X(t) = x) &= \alpha_j(x)dt + o(dt), \\ \mathbb{P}(R_j(t+dt) - R_j(t) > 1 | X(t) = x) &= o(dt), \end{aligned}$$

for  $dt \rightarrow 0$  and  $x \in \mathbb{X}$ .

An immediate question is the existence of a set of stochastic processes  $\{R_1, \dots, R_m\}$  that satisfy Assumption 2.19. Consider the reaction jump process

$$X : [0, \infty) \times \Omega \rightarrow \mathbb{N}_0^m$$

together with his natural filtration  $\{\mathcal{F}_t\}_{t \in [0, \infty)}$ ,  $\mathcal{F}_t = \sigma(X(s) : s \leq t)$ , and probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  where  $\bigvee_{0 \leq t} \mathcal{F}_t = \mathcal{F}$ . Recall that the function

$$s \mapsto X(s, \omega)$$

is right continuous for  $\mathbb{P}$ -a.s.  $\omega$ , and since  $(\mathbb{N}_0^m, \rho)$  ( $\rho$  is the discrete metric) is a separable metric space, it follows from Theorem 3 in [10] that for all  $t > 0$ ,  $X : [0, t] \times \Omega \rightarrow \mathbb{N}_0^m$  is  $\mathbb{B}([0, t]) \times \mathcal{F}_t$ -measurable. Then if  $\alpha_j : \mathbb{X} \rightarrow \mathbb{R}_0^+$  is the propensity function of  $\mathcal{R}_j$ ,

$$(s, \omega) \mapsto \alpha_j(X(s, \omega))$$

is  $\mathbb{B}([0, t]) \times \mathcal{F}_t$ -measurable and since is a non-negative function, it follows from Tonelli theorem that

$$\omega \mapsto f_j(t, \omega) := \int_0^t \alpha_j(X(s, \omega)) ds$$

is also  $\mathcal{F}_t$ -measurable.

Observe that if  $\mathbb{P}(f_j(t+h) - f_j(t) > 0) > 0$ , then

$$\mathbb{P}(\omega \in \Omega : \alpha_j(X(s, \omega)) > 0 \quad \forall s \in I(\omega) \quad \text{for some subinterval } I(\omega) \subset (t, t+h]) > 0,$$

which implies that with positive probability a new reaction  $\mathcal{R}_j$  could happen in the time interval  $(t, t+h]$ . That is, the stochastic process  $\{f_j(t)\}$  give us the cumulative propensity of  $\mathcal{R}_j$ , but it doesn't count the number of reactions  $\mathcal{R}_j$  through time. In order to do this counting consider a unit-rate Poisson process  $\{\mathcal{U}_j(t)\}$  adapted to filtration  $\{\mathcal{F}_t\}$ ; recall that this process can be written as

$$\mathcal{U}_j(t) = \max\{k \in \mathbb{N}_0 : S_{j,k} \leq t\}$$

where

$$S_{j,0} := 0, \quad S_{j,k} := \sum_{i=1}^k s_{j,i}$$

and  $\{s_{j,k}\}_k$  is an i.i.d. sequence with distribution  $\exp(1)$ . This implies that the average time between events is 1. With this amount of information let

$$(2.20) \quad R_j : [0, \infty) \times \Omega \rightarrow \mathbb{N}_0 \quad R_j(t, \omega) := \mathcal{U}_j(f_j(t, \omega), \omega)$$

First of all, observe that for a.s.  $\omega \in \Omega$   $t \mapsto R_j(t, \omega)$  is right continuous, because for a.s.  $\omega \in \Omega$ ,  $t \mapsto f_j(t, \omega)$  is an increasing process and for a.s.  $\omega \in \Omega$ ,  $t \mapsto \mathcal{U}_j(t, \omega)$  is right continuous. Second of all, for each  $t \in [0, \infty)$  what can we say about the function  $\omega \mapsto R_j(t, \omega)$ ? Is  $\omega \mapsto R_j(t, \omega)$  a random variable for each  $t \in [0, \infty)$ ; and if it is a random variable, is it  $\mathcal{F}_t$ -measurable?; that is, is  $\{R_j(t)\}_t$  a stochastic process adapted to filtration  $\{\mathcal{F}_t\}$ ? And finally, since we are working with unit-rate Poisson process  $\mathcal{U}_j$  that satisfies

$$E[\mathcal{U}_j(t)] = t = E[t]$$

what can we say about  $E[\mathcal{U}_j(f_j(t))]$ ? Is it true that  $E[\mathcal{U}_j(f_j(t))] = E[f_j(t)]$ ?

The answers for these questions are positive; see for example [8].

Now we are going to show that this definition of  $R_j(t)$  satisfies the Fundamental Hypothesis. By the properties of HPP, we have

$$\mathcal{U}_j(f_j(t + dt)) = \mathcal{U}_j((f_j(t), f_j(t + dt))) + \mathcal{U}_j(f_j(t)).$$

Hence

$$R_j(t + dt) = \mathcal{U}_j((f_j(t), f_j(t + dt))) + R_j(t)$$

and

$$R_j(t + dt) - R_j(t) = \mathcal{U}_j((f_j(t), f_j(t + dt))).$$

Since the propensity functions are constant in the infinitesimal time interval  $[t, t + dt)$ , we get

$$\begin{aligned} \mathbb{P}(R_j(t + dt) - R_j(t) = 1 | X(t) = x) &= \mathbb{P}(\mathcal{U}_j((f_j(t), f_j(t + dt))) = 1 | X(t) = x) \\ &= \mathbb{P}(\mathcal{U}_j(\alpha_j(x) dt) = 1) \\ &= \exp(-\alpha_j(x) dt) \\ &= \alpha_j(x) dt + o(dt). \end{aligned}$$

Therefore (2.18) can be written as

$$(2.21) \quad X(t) = x_0 + \sum_{j=1}^n \mathcal{U}_j \left( \int_0^t \alpha_j(X(s)) ds \right) v[j],$$

which is the Anderson-Kurtz's Equation for this chemical system. In (2.21)

$$\{\mathcal{U}_1, \dots, \mathcal{U}_n\}$$

is a set of independent unit-rate Poisson processes with sequence of times events  $\{S_{j,k}\}_{k \in \mathbb{N}}$  for each  $\mathcal{U}_j$ . From the theory of HPP we know that for  $i \neq j$ ,

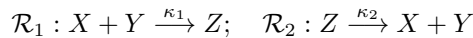
$$S_{j,k} \neq S_{i,k} \quad \forall k \in \mathbb{N},$$

hence every solution  $\{X(t)\}$  of (2.21) makes jumps, with probability 1, of the form  $x \rightarrow x + v[j]$  whenever one of the Poisson processes  $\mathcal{U}_j$  indicates such a jump. That is, if  $\{X(t)\}$  is a solution of (2.21), then  $\{X(t)\}$  must satisfy

$$X(t) = \sum_{k=0}^{\infty} x_k \chi_{[t_k, t_{k+1})}(t)$$

where  $\forall k \in \mathbb{N}_0$   $x_{k+1} = x_k + w$  for some  $w \in \{v[1], \dots, v[n]\}$  and  $\{t_k\}$  are the times when a stoichiometric change occurs due to a chemical reaction. In Chapter 4 we will show that (2.21) has as unique solution the reaction jump process given in Section 2.2.

**Example 2.22.** (*Binding and Unbinding*) For the reaction system of binding and unbinding



the path representation  $X(t) = (x(t), y(t), z(t))^T$  is given by

$$\begin{aligned} x(t) &= x(0) - R_1(t) + R_2(t) \\ y(t) &= y(0) - R_1(t) + R_2(t) \\ z(t) &= z(0) + R_1(t) - R_2(t) \end{aligned}$$

where  $x(t)$ ,  $y(t)$  and  $z(t)$  are the number of molecules of  $X$ ,  $Y$  and  $Z$ , respectively, at time  $t$ . Since  $\mathcal{R}_1$  has propensity  $\alpha_1(x, y, z) = \kappa_1 xy$  and  $\mathcal{R}_2$  has propensity  $\alpha_2(x, y, z) = \kappa_2 z$  we can write  $\{X(t)\}$  as

$$X(t) = X(0) + \mathcal{U}_1(\kappa_1 \int_0^t x(s)y(s)ds) \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} + \mathcal{U}_2(\kappa_2 \int_0^t z(s)ds) \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

In this equation, the second term on the right-hand side stands the number of times that the reaction  $\mathcal{R}_1$  happens in the time interval  $[0, t]$  and the last term stands the number of times that reaction  $\mathcal{R}_2$  happens in the time interval  $[0, t]$ .

**2.4. Backward and Forward Kolmogorov's Equations.** So far we have seen how Anderson-Kurtz's equation can help us to describe the dynamics of a chemical reaction system; nevertheless, there is another popular and alternative form to describe the time dynamics of a chemical reaction system; the Forward Kolmogorov's Equation, also called Chemical Master Equation in the biochemical field, see for example [9].

Given a conservative and stable infinitesimal generator, say  $A$ , with space of states

$\mathbb{X}$ , we are interested in the transition probabilities between states in  $\mathbb{X}$  in a certain period of time. That is, if  $\{X(t)\}$  is the HMC defined by  $A$ , we are looking for the values

$$p_{i,j}(h) := \mathbb{P}(X(t+h) = j | X(t) = i) \quad \forall i, j \in \mathbb{X}.$$

Let  $P(t) = [p_{i,j}(t)]_{i,j \in \mathbb{X}} \forall t \geq 0$ . The fact that  $\{X(t)\}$  is a continuous-time HMC implies that  $\{P(t)\}$  satisfies

$$\forall i, j \in \mathbb{X}, \quad p_{i,j}(t) \geq 0, \quad \sum_{j \in \mathbb{X}} p_{i,j}(t) = 1$$

and the Chapman-Kolmogorov Equations:

$$(2.23) \quad \forall i, j \in \mathbb{X}, p_{i,j}(t+h) = \sum_{k \in \mathbb{X}} p_{i,k}(t) p_{k,j}(h),$$

and

$$(2.24) \quad \forall i, j \in \mathbb{X}, p_{i,j}(t+h) = \sum_{k \in \mathbb{X}} p_{i,k}(h) p_{k,j}(t).$$

Following the standard notation in the theory of continuous time HMC, we define  $q_i := -A_{i,i}$  and  $q_{i,j} := A_{i,j}$ . Therefore,  $\{P(t)\}$  satisfies (see for example [6])

$$(2.25) \quad \forall i, j \in \mathbb{X}, \quad q_i = \lim_{h \downarrow 0} \frac{1 - p_{i,i}(h)}{h}, \quad q_{i,j} = \lim_{h \downarrow 0} \frac{p_{i,j}(h)}{h}.$$

Using Fatou's lemma  $\sum_{j \neq i} q_{i,j}$  and  $q_i$  satisfy

$$\begin{aligned} \sum_{j \neq i} q_{i,j} &= \sum_{j \neq i} \liminf_{h \downarrow 0} \frac{p_{i,j}(h)}{h} \\ &\leq \liminf_{h \downarrow 0} \sum_{j \neq i} \frac{p_{i,j}(h)}{h} \\ &= \liminf_{h \downarrow 0} \frac{1 - p_{i,i}(h)}{h} \\ &= q_i. \end{aligned}$$

That is if  $A$  is a generator, not necessarily conservative, then  $\sum_{j \neq j} q_{i,j}$  and  $q_i$  always satisfy

$$(2.26) \quad \sum_{j \neq i} q_{i,j} \leq q_i.$$

Observe that (2.24) implies

$$\begin{aligned} p_{i,j}(t+h) - p_{i,j}(t) &= \sum_{k \in \mathbb{X}} p_{i,k}(h) p_{k,j}(t) - p_{i,j}(t) \\ &= p_{i,j}(t)(p_{i,i}(h) - 1) + \sum_{k \neq i} p_{i,k}(h) p_{k,j}(t) \end{aligned}$$

and

$$(2.27) \quad \frac{p_{i,j}(t+h) - p_{i,j}(t)}{h} = \frac{p_{i,i}(h) - 1}{h} p_{i,j}(t) + \sum_{k \neq i} \frac{p_{i,k}(h)}{h} p_{k,j}(t)$$

Using equation (2.27) and Fatou's lemma we get

$$\liminf_{h \rightarrow 0} \frac{p_{i,j}(t+h) - p_{i,j}(t)}{h} \geq -q_i p_{i,j}(t) + \sum_{k \neq i} q_{i,k} p_{k,j}(t).$$

Now, since  $A$  is stable and conservative, we get for  $N > i$

$$\frac{1}{h} \sum_{j>N} p_{i,j}(h) = \frac{1}{h} (1 - \sum_{j \leq N} p_{i,j}(h)) \rightarrow q_i - \sum_{j<N, j \neq i} q_{i,j} = \sum_{j>N} q_{i,j}.$$

Note that from the conservative property, the convergence of the series  $\frac{1}{h} \sum_{j>N} p_{i,j}(h)$  only depends on the convergence of the finite series  $\frac{1}{h} (1 - \sum_{j \leq N} p_{i,j}(h))$  which converges uniformly in  $h$ . Hence  $\sum_{j \neq i} \frac{p_{i,j}(h)}{h}$  converges uniformly to  $\sum_{j \neq i} q_{i,j}$ . Therefore, if we let  $h \rightarrow 0$  in (2.27), we get the Backward Kolmogorov Equations

$$(2.28) \quad p_{i,j}^{(1)}(t) = -q_i p_{i,j}(t) + \sum_{k \neq i} q_{i,k} p_{k,j}(t).$$

This proves an implication of the following Theorem(see [7]).

**Theorem 2.29.** *A stable infinitesimal generator  $A$  is conservative if and only if its transition semigroup  $\{P(t)\}$  satisfies the Backward Kolmogorov Equations, that is*

$$P^{(1)}(t) = AP(t).$$

*Proof.* We only need to prove the if part. Let  $A$  be a stable generator such that his semigroup satisfies the Backward Kolmogorov equations (2.28). Then, for fixed  $i \in \mathbb{X}$  we get  $\forall t \geq 0$

$$\begin{aligned} \sum_{j \in \mathbb{X}} p_{i,j}^{(1)}(t) &= \sum_{j \in \mathbb{X}} -q_i p_{i,j}(t) + \sum_{j \in \mathbb{X}} \sum_{k \neq i} q_{i,j} p_{k,j}(t) \\ &= -q_i \sum_{j \in \mathbb{X}} p_{i,j}(t) + \sum_{k \neq i} q_{i,j} \sum_{j \in \mathbb{X}} p_{k,j}(t) \\ &= -q_i + \sum_{k \neq i} q_{i,j} \end{aligned}$$

and, using inequality(2.26),

$$\begin{aligned} |\sum_{j=1}^n p_{i,j}^{(1)}(t)| &= | -q_i \sum_{j=1}^n p_{i,j}(t) + \sum_{k \neq i} q_{i,k} (\sum_{j=1}^n p_{k,j}(t)) | \\ &\leq q_i \sum_{j=1}^n p_{i,j}(t) + \sum_{k \neq i} q_{i,k} (\sum_{j=1}^n p_{k,j}(t)) \\ &\leq q_i + \sum_{k \neq i} q_{i,k} \leq 2q_i. \end{aligned}$$

It follows using dominated convergence theorem that

$$\begin{aligned} \int_{[s, s+\delta]} \sum_{j \in \mathbb{X}} p_{i,j}^{(1)}(t) dt &= \sum_{j \in \mathbb{X}} \int_{[s, s+\delta]} p_{i,j}^{(1)}(t) dt \\ &= \sum_{j \in \mathbb{X}} [p_{i,j}(s+\delta) - p_{i,j}(s)] \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^n [p_{i,j}(s+\delta) - p_{i,j}(s)] \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^n p_{i,j}(s+\delta) - \sum_{j=1}^n p_{i,j}(s) \\ &= 1 - 1 = 0. \end{aligned}$$

Since  $\sum_{j \in \mathbb{X}} p_{i,j}^{(1)}(t) = -q_i + \sum_{k \neq i} q_{i,j} \forall t \geq 0$ , we get

$$-q_i + \sum_{j \neq i} q_{i,j} = 0.$$

□

Backward Kolmogorov's Equations have an important probabilistic interpretation. Suppose that at time  $X(t_0) = i$ . Then the probability that  $X(t_0 + t) = k$  and the first discontinuity is a jump in the time interval  $(t_0, t_0 + t]$ , is given by

$$\mathbb{P}(X(t_0 + t) = k, t_0 < \tau_1 < t_0 + t | X(t_0) = i) = \int_0^t q_i \exp(-q_i s) \frac{q_{i,j}}{q_i} p_{j,k}(t-s) ds,$$

see for example [7]. Therefore,  $\forall i, k \in \mathbb{X}$

$$(2.30) \quad p_{i,k}(t) \geq \sum_{j \neq i} \int_0^t \exp(-q_i(t-s)) q_{i,j} p_{j,k}(s) ds + \delta_{i,k} \exp(-q_i t).$$

Since the equality in (2.30) is an integral form of the equation (2.28), if the semi-group  $\{P(t)\}$  satisfies the Backward Kolmogorov Equations, then  $\forall t_0 > 0$  with probability 1, the first discontinuity of  $X(t)$  following  $t_0$  is a jump.

With the help of (2.23) we get

$$\begin{aligned} p_{i,j}(t+h) - p_{i,j}(t) &= \sum_{k \in \mathbb{X}} p_{i,k}(t) p_{k,j}(h) - p_{i,j}(t) \\ &= p_{i,j}(t)(p_{j,j}(h) - 1) + \sum_{k \neq j} p_{i,k}(t) p_{k,j}(h) \end{aligned}$$

and

$$(2.31) \quad \frac{p_{i,j}(t+h) - p_{i,j}(t)}{h} = p_{i,j}(t) \frac{p_{j,j}(h) - 1}{h} + \sum_{k \neq j} p_{i,k}(t) \frac{p_{k,j}(h)}{h}.$$

It follows from (2.31) and Fatou's lemma

$$\liminf_{h \rightarrow 0} \frac{p_{i,j}(t+h) - p_{i,j}(t)}{h} \geq -p_{i,j}(t) q_j + \sum_{k \neq j} p_{i,k}(t) q_{k,j}.$$

We have seen that  $\forall t > 0$  and  $i, j \in \mathbb{X}$  the term  $p_{i,j}^{(1)}(t)$  exists when the generator  $A$  is stable and conservative. Therefore

$$(2.32) \quad p_{i,j}^{(1)}(t) \geq -p_{i,j}(t) q_j + \sum_{k \neq j} p_{i,k}(t) q_{k,j}$$

The following theorem gives us conditions under which there is equality in (2.32). For more details of the proof see [4].

**Theorem 2.33.** *Let  $A$  be a conservative and stable generator. Suppose that  $\forall i \in \mathbb{X}$  and  $\forall t \geq 0$*

$$\sum_{k \in \mathbb{X}} p_{i,k}(t) q_k < \infty.$$

*Then the Forward Kolmogorov equations*

$$(2.34) \quad p_{i,j}^{(1)}(t) = -p_{i,j}(t) q_j + \sum_{k \neq j} p_{i,k}(t) q_{k,j}$$

*are satisfied.*

*Proof.*

$$\frac{p_{k,j}(h)}{h} \leq \frac{1 - p_{k,k}(h)}{h} \leq q_k$$

Hence,  $\forall t > 0$

$$p_{i,k}(t) \frac{p_{k,j}(h)}{h} \leq p_{i,k}(t) q_k.$$

Since  $\sum_{k \neq j} p_{i,k}(t) q_k < \infty$ , by the dominated convergence theorem

$$\lim_{h \rightarrow 0} \sum_{k \neq j} p_{i,k}(t) \frac{p_{k,j}(h)}{h} = \sum_{k \neq j} p_{i,k}(t) \lim_{h \rightarrow 0} \frac{p_{k,j}(h)}{h} = \sum_{k \neq j} p_{i,k}(t) q_{k,j}.$$

Using (2.31) and the above equality we get the Forward Kolmogorov equation.  $\square$

Now we are going to find an integral form for a semigroup (with stable and conservative generator) that satisfies the Forward Kolmogorov Equations and the strong hypothesis

$$\sum_{k \in \mathbb{X}} q_k < \infty.$$

Let

$$b(t) = \sum_{k \neq j} p_{i,k}(t) q_{k,j}.$$

Observe that  $\forall t > 0$ , the product  $p_{i,k}(t) q_{k,j} \leq q_{k,j} \leq q_k$ . Since  $\sum_{k \neq j} q_k < \infty$ , it follows by the dominated convergence theorem and the continuity of the semigroup

$$\lim_{h \rightarrow 0} b(t+h) = \sum_{k \neq j} \lim_{h \rightarrow 0} p_{i,k}(t+h) q_{k,j} = b(t),$$

i.e.,  $b$  is a continuous function.

We can write the Forward Kolmogorov equations as follows

$$p_{i,j}^{(1)}(t) = -q_j p_{i,j}(t) + b(t),$$

and by the variation of parameters formula

$$p_{i,j}(t) = \delta_{i,j} \exp(-q_i t) + \int_0^t \left( \sum_{k \neq j} p_{i,k}(t-s) q_{k,j} \right) \exp(-q_j s) ds.$$

If  $\phi : \mathbb{R}_0^+ \times (\mathbb{N} \setminus \{j\}) \rightarrow [0, 1]$  is given by

$$\phi(t, k) = p_{i,k}(t),$$

then  $\phi^{-1}((a, b]) = \bigcup_{k \in (\mathbb{N} \setminus \{j\})} p_{i,k}^{-1}((a, b]) \times \{k\} \in \mathbb{B}(\mathbb{R}_0^+) \times \mathcal{P}(\mathbb{N} \setminus \{j\})$ , because  $\forall k \in \mathbb{N} \setminus \{j\}$  the term  $p_{i,k}$  is a continuous function. Using Tonelli's theorem we get

$$(2.35) \quad p_{i,j}(t) = \delta_{i,j} \exp(-q_i t) + \sum_{k \neq j} \int_0^t p_{i,k}(t-s) q_{k,j} \exp(-q_j s) ds.$$

As Doob showed in [7], the probabilistic meaning of (2.35) is that  $\forall t > 0$ , the last discontinuity of the continuous-time HMC is a jump with probability 1. Therefore, the semigroup satisfies the Forward Kolmogorov equations if and only if  $\forall t > 0$  the last discontinuity of the associated continuous-time HMC is a jump with probability 1. Hence, if the semigroup satisfies the Forward Kolmogorov equations,

the associated continuous-time HMC is regular.

In the following section, we are going to show Gillespie's proof that the semigroup of the reaction jump process satisfies the Chemical Master Equation, which is the name given to the corresponding Forward Kolmogorov Equation. This ensures that every reaction jump process must be regular, whether  $\mathbb{X}$  is finite or infinite.

**2.5. Chemical Master Equation(CME).** We start this section analysing the Forward Kolmogorov Equations, also known as Chemical Master Equation, for the reaction system of  $n$  reactions and  $m$  species. Since we are supposing that the space of states  $\mathbb{X} \subset \mathbb{N}^m$  is finite, Theorem 2.33 implies that the chemical system satisfies these equations. Let  $x_0, x \in \mathbb{X}$  and define

$$P(x, t|x_0, t_0) := \mathbb{P}(X(t) = x|X(t_0) = x_0).$$

Then, the CME is given by

$$\frac{dP(x, t|x_0, t_0)}{dt} = A(x, x)P(x, t|x_0, t_0) + \sum_{y \neq x} A(y, x)P(y, t|x_0, t_0),$$

where  $A$  is given by (2.11), (2.12) and (2.13). Therefore

$$(2.36) \quad \frac{dP(x, t|x_0, t_0)}{dt} = \sum_{j=1}^n [\alpha_j(x - v[j])P(x - v[j], t|x_0, t_0) - \alpha_j(x)P(x, t|x_0, t_0)].$$

Observe that the previous derivation of the CME lacks a proper chemical interpretation, which is a problem when we are working in applications. In order to solve this problem, we are going to show the proof, given by Gillespie, that the reaction system satisfies (2.36); suppose that at time  $t_0$  the system is in  $x_0$ . Let

$$\mathbb{P}_{x_0, t_0}(\ast) := \mathbb{P}(\ast|X(t_0) = x_0)$$

be the conditional probability to the event  $\{X(t_0) = x_0\}$  and let  $E$  be the event

$$E := \{X(t + dt) = x\},$$

where  $(t, t + dt]$  such that the probability that more than one reaction takes place in this interval is  $o(dt)$  (See Assumption 2.19). Thus we can take  $dt$  such that at most one reaction can take place over  $(t, t + dt]$ . Now, notice that to be in the state  $x$  at time  $t + dt$ , there are only two basic scenarios for time  $t$ ; either the system was already in the state  $x$  at time  $t$  and no reaction took place over  $(t, t + dt]$ , or for some  $1 \leq j \leq n$  the system was in the state  $x - v[j]$  at time  $t$  and the  $j$ th reaction fired over  $[t, t + dt)$ , thereby bringing the system into state  $x$ . Define

$$H_0 := \{X(t) = x\},$$

for  $1 \leq j \leq n$

$$H_j := \{X(t) = x - v[j]\}$$



and

$$H_{n+1} := \{X(t) = y : y \neq x - v[j], \quad 1 \leq j \leq n\}.$$

Since  $\{H_0, H_1, \dots, H_n, H_{n+1}\}$  is a partition of the probability space, then

$$P(x, t + dt | x_0, t_0) = \mathbb{P}_{x_0, t_0}(E) = \sum_{j=0}^{n+1} \mathbb{P}_{x_0, t_0}(E | H_j) \mathbb{P}_{x_0, t_0}(H_j),$$

where  $\mathbb{P}_{x_0, t_0}(H_0) = P(x, t | x_0, t_0)$ ,  $\mathbb{P}_{x_0, t_0}(H_j) = P(x - v[j], t | x_0, t_0)$  with  $1 \leq j \leq n$ . By the Markov property of the reaction chain  $\{X(t)\}$  and the fundamental hypothesis, for  $1 \leq j \leq n$

$$\begin{aligned} \mathbb{P}_{x_0, t_0}(E | H_j) &= P(x, t + dt | x - v[j], t) = \alpha_j(x - v[j])dt, \\ \mathbb{P}_{x_0, t_0}(E | H_0) &= P(x, t + dt | x, t) = 1 - \left( \sum_{j=1}^n \alpha_j(x) \right) dt + o(dt) \end{aligned}$$

and

$$\mathbb{P}_{x_0, t_0}(E | H_{n+1}) = \mathbb{P}(X(t + dt) = x | \{X(t) = y : y \neq x - v[j], \quad 1 \leq j \leq n\}) = o(dt).$$

Therefore

$$\begin{aligned} &\frac{P(x, t + dt | x_0, t_0) - P(x, t | x_0, t_0)}{dt} \\ &= \left( \sum_{j=1}^n \alpha_j(x - v[j]) P(x - v[j], t | x_0, t_0) \right) - \left( \sum_{j=1}^n \alpha_j(x) P(x, t | x_0, t_0) \right) + \frac{o(dt)}{dt} \end{aligned}$$

and

$$\begin{aligned} &\lim_{dt \downarrow 0} \frac{P(x, t + dt | x_0, t_0) - P(x, t | x_0, t_0)}{dt} \\ &= \sum_{j=1}^n [\alpha_j(x - v[j]) P(x - v[j], t | x_0, t_0)] - \alpha_j(x) P(x, t | x_0, t_0). \end{aligned}$$

Note that  $\frac{dP(x, t | x_0, t_0)}{dt}$  exists because the generator  $A$ , of the reaction chain, is stable and conservative. Then

$$\frac{dP(x, t | x_0, t_0)}{dt} = \sum_{j=1}^n [\alpha_j(x - v[j]) P(x - v[j], t | x_0, t_0)] - \alpha_j(x) P(x, t | x_0, t_0).$$

Observe that this proof of the CME only uses the Fundamental Hypothesis. Thus, in the hypothetical case where the space of states of the reaction system is infinite, the corresponding semigroup satisfies the Forward Kolmogorov Equations which implies that the reaction process is a regular continuous-time HMC.

Define  $\pi_x(t) = \mathbb{P}(X(t) = x)$  for  $t \in [0, \infty)$  and  $x \in \mathbb{X}$ . Consider an initial distribution  $\pi_x(0) = \mathbb{P}(X(0) = x)$  for the reaction jump process. Then,  $\forall t \geq 0$   $\pi(t)^T = \pi(0)^T P(t)$  and from the CME  $\forall x \in \mathbb{X}$

$$(2.37) \quad \frac{d}{dt} \pi_x(t) = \sum_{j=1}^n [\alpha_j(x - v[j]) \pi_{x-v[j]}(t) - \alpha_j(x) \pi_x(t)],$$

or in matrix representation

$$(2.38) \quad \frac{d}{dt} \pi(t) = A^T \pi(t).$$

**Proposition 2.39.**  $\pi$  is a stationary signed measure of  $\{P(t)\}$  if and only if  $\pi$  is an equilibrium point of (2.38).

*Proof.* Suppose that  $\pi$  is a stationary signed measure of  $\{P(t)\}$ , that is  $\pi = \pi^T P(t)$ . Then,  $0 = \pi^T P'(t) = \pi^T AP(t)$ . If  $t \downarrow 0$ , then  $0 = \pi^T A$ .

Let  $\pi$  be an equilibrium point of (2.38) and define  $\forall t \geq 0$   $\pi(t)^T = \pi^T P(t)$ ; then,

$$\frac{d}{dt}\pi(t)^T = \pi^T P'(t) = \pi^T AP(t) = 0.$$

Since  $P(0) = I$ ,  $\pi = \pi^T P(t)$ .  $\square$

For obvious reasons, we are interested in those equilibrium points that are positive. Zeron and Santillan have shown that (2.38) have at least one stable nonzero stationary probability distribution and, whenever the reaction jump process is irreducible, this probability distribution is unique and asymptotically stable, see for example [19]. Now consider the embedded chain  $\{X_n\}$  of an irreducible reaction jump process; if  $\pi$  is the only stationary probability distribution of the process, then  $\mu$  defined  $\forall x \in \mathbb{X}$  as

$$\mu(x) := \frac{\pi(x)\alpha_{sum}(x)}{C}$$

is a stationary distribution of the embedded chain, where  $C = \sum_{x \in \mathbb{X}} \pi(x)\alpha_{sum}(x)$ . Moreover, the following proposition implies that is unique  $\mu$ . This result can be found in [4].

**Proposition 2.40.** Let  $\{X_n\}$  be an irreducible HMC with transition matrix  $P$ . If  $\{X_n\}$  has a stationary probability distribution, then  $\{X_n\}$  is recurrent.

*Proof.* Let  $\pi$  be a stationary probability distribution ( $\pi^T = \pi^T P$ ). Then,  $\forall n \in \mathbb{N}$   $\pi^T = \pi^T P^n$  and for all state of the chain  $j$ ,  $\pi(j) = \sum_i \pi(i)p_{i,j}(n)$ . If some state  $j$  is transient, then  $\sum_n \chi_{\{X_n=j\}} < \infty$  a.s. It follows that  $\lim_{n \rightarrow \infty} \chi_{\{X_n=j\}} = 0$  a.s. and by the dominated convergence theorem ( $\chi_{\{X_n=j\}} \leq 1$ )

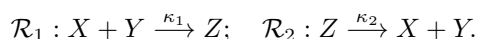
$$\lim_{n \rightarrow \infty} p_{i,j}(n) = \lim_{n \rightarrow \infty} E_i[\chi_{\{X_n=j\}}] = E_i[\lim_{n \rightarrow \infty} \chi_{\{X_n=j\}}] = 0.$$

Again by the dominated convergence theorem ( $\pi(i)p_{i,j}(n) \leq \pi(i)$  and  $\sum_i \pi(i) = 1$ )

$$\pi(j) = \lim_{n \rightarrow \infty} \sum_i \pi(i)p_{i,j}(n) = \sum_i \pi(i) \lim_{n \rightarrow \infty} p_{i,j}(n) = 0.$$

Since the chain is irreducible every state is transient which implies that  $\pi = 0$  and this can not happen  $\square$

**Example 2.41.** (Irreducible reaction jump process) Consider the reaction system of binding and unbinding



Suppose that at time 0 the system is in the state  $(x_0, y_0, 0)^T$ , where  $x_0$  and  $y_0$  are the number of molecules of  $X$  and  $Y$ , respectively. Then the space of states is given

by

$$\mathbb{X} = \{(x_0 - b, y_0 - b, b)^T : 0 \leq b \leq \min(x_0, y_0)\}.$$

Let  $0 \leq c < b \leq \min(x_0, y_0)$  and  $\hat{x}_b = (x_0 - b, y_0 - b, b)^T$ ,  $\hat{x}_c = (x_0 - c, y_0 - c, c)^T$ . Observe  $\hat{x}_c = \hat{x}_b + (b - c)v[2]$  and  $\hat{x}_b = \hat{x}_c + (b - c)v[1]$ ; that is we can go from  $\hat{x}_b$  to  $\hat{x}_c$  in  $b - c$  reactions of type  $\mathcal{R}_2$  and from  $\hat{x}_c$  to  $\hat{x}_b$  in  $b - c$  reactions of type  $\mathcal{R}_1$ :

$$\begin{aligned} \hat{x}_b &\xrightarrow{\mathcal{R}_2} \hat{x}_b + v[2] \xrightarrow{\mathcal{R}_2} \dots \xrightarrow{\mathcal{R}_2} \hat{x}_b + (b - c - 1)v[2] \xrightarrow{\mathcal{R}_2} \hat{x}_c, \\ \hat{x}_c &\xrightarrow{\mathcal{R}_1} \hat{x}_c + v[1] \xrightarrow{\mathcal{R}_1} \dots \xrightarrow{\mathcal{R}_1} \hat{x}_c + (b - c - 1)v[1] \xrightarrow{\mathcal{R}_1} \hat{x}_b. \end{aligned}$$

If  $\{X_n\}$  is the embedded HMC of the reaction jump process, then

$$\begin{aligned} &\mathbb{P}(X_{b-c} = \hat{x}_c | X_0 = \hat{x}_b) \\ &\geq \mathbb{P}(X_1 = \hat{x}_b + v[2], \dots, X_{b-c-1} = \hat{x}_b + (b - c - 1)v[2], X_{b-c} = \hat{x}_c | X_0 = \hat{x}_b) \\ &= \mathbb{P}(X_1 = \hat{x}_b + v[2] | X_0 = \hat{x}_b) \dots \mathbb{P}(X_{b-c} = \hat{x}_c | X_{b-c-1} = \hat{x}_b + (b - c - 1)v[2]) \\ &= \prod_{k=0}^{b-c-1} \frac{\alpha_2(\hat{x}_b + kv[2])}{\alpha_{sum}(\hat{x}_b + kv[2])} > 0, \end{aligned}$$

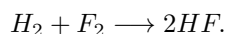
and

$$\mathbb{P}(X_{b-c} = \hat{x}_b | X_0 = \hat{x}_c) = \prod_{k=0}^{b-c-1} \frac{\alpha_2(\hat{x}_c + kv[1])}{\alpha_{sum}(\hat{x}_c + kv[1])} > 0.$$

Thus the embedded chain is irreducible, which implies that the jump process is irreducible.

### 3. CHEMICALS REACTIONS WITH DELAYS

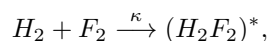
In Chapter 2 we introduced and modeled the chemical reaction where one molecule of gaseous hydrogen and one molecule of gaseous fluorine react to form two molecules of hydrogen fluoride:



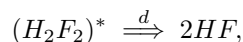
Although, for practical reasons, this reaction is thought to happen instantaneously, the reality is that the molecules of  $H_2$  and  $F_2$  firstly form an excited complex, denoted by  $(H_2F_2)^*$ , and then this complex decomposes into two molecules of  $HF$  after a short period of time (see for example [18] and [11]). Therefore, the reaction of gaseous hydrogen and gaseous fluorine can be written more properly as follows



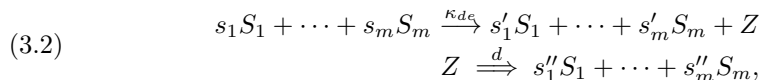
where  $\xrightarrow{d}$  denotes a physical transition that takes  $d$  units of time. Together the chemical reaction



and the physical transition



are called a chemical reaction with delay; and this will be our general scheme. Thus, consider a chemical system of  $m$  chemical species  $S_1, \dots, S_m$  and space of states  $\mathbb{X} \subset \mathbb{N}_0^m$ ; then a chemical reaction with delay  $\mathcal{R}_{de}$ , that produces an excited complex  $Z$ , is represented by stoichiometric equations of the form:



with stoichiometric coefficients  $s_j$ ,  $s'_j$  and  $s''_j$  in  $\mathbb{N}_0$ . Note that in (3.2) is implicit the fact that we are making a distinction between the chemical species  $S_1, \dots, S_m$  and the excited complex  $Z$ . Now  $\mathcal{R}_{de}$  introduces two stoichiometric vectors given by

$$\omega_1 := (s'_1 - s_1, \dots, s'_m - s_m)^T$$

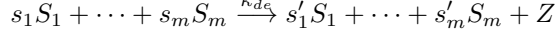
and

$$\omega_2 := (s''_1, \dots, s''_m)^T.$$

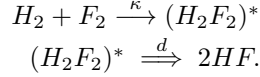
Therefore, following the ideas presented in Section 2.1 we can write  $\mathcal{R}_{de}$  as

$$\mathcal{R}_{de} : \begin{cases} x \xrightarrow{\kappa_{de}} x + \omega_1 \\ Z \xrightarrow{d} \omega_2 \end{cases}$$

for each  $x \in \mathbb{X}$  such that  $x + \omega_1 \in \mathbb{X}$ . In the same way that every chemical reaction has its own propensity function, the chemical reaction part of  $\mathcal{R}_{de}$



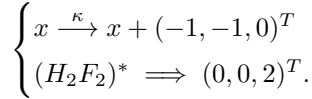
also has his own propensity function  $\alpha_{de}$ , which satisfies (2.5) and (2.7). Hence,  $\mathcal{R}_{de}$  is characterized by four parameters: the time of the physical transition  $d$ , the two stoichiometric vectors  $\omega_1, \omega_2$  and the propensity function  $\alpha_{de}$  of the chemical reaction part of  $\mathcal{R}_{de}$ . For example, consider again the delay reaction



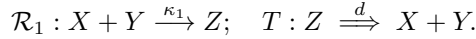
In this example  $\omega_1 = (-1, -1, 0)^T$ ,  $\omega_2 = (0, 0, 2)^T$  and

$$\alpha(x_1, x_2, x_3) = \kappa x_1 x_2,$$

where  $x_1, x_2$  and  $x_3$  are the number of molecules of  $H_2, F_2$  and  $HF$ , respectively. Hence, this reaction can be written as



**Example 3.3.** (*Binding and Unbinding Reaction with Delay*) Consider the reaction system with chemical species  $X, Y$  and one complex  $Z$ ; moreover, add a chemical reaction with delay composed by a normal  $\mathcal{R}_1$  and a physical transition  $T$  with delay  $d$



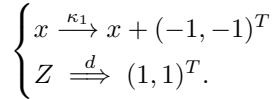
Together,  $\mathcal{R}_1$  and  $T$  constitute a Reaction with Delay

$$\mathcal{R}_{de} = \mathcal{R}_{de}(d, \omega_1, \omega_2, \alpha_{de}),$$

where  $\omega_1 = (-1, -1)$ ,  $\omega_2 = (1, 1)$ , and  $\alpha_{de} : \mathbb{X} \rightarrow \mathbb{R}_+$  is given by  $\alpha_{de}(x, y) = \kappa_1 xy$ . If the initial number of molecules of  $X$  and  $Y$  are  $x_0$  and  $y_0$  ( $x_0 < y_0$ ), respectively, then

$$\mathbb{X} = \{(x_0, y_0)^T, (x_0 - 1, y_0 - 1), \dots, (0, y_0 - x_0)^T\}.$$

Hence



**3.1. Reaction Jump Process of a Chemical System with one Delayed Reaction.** Now that we have an interpretation of what a chemical reaction with delay is, and we also have a proper form to model it, the following question is: how can we represent the dynamics of a chemical system with one delayed reaction? As we saw in Chapter 2, the reaction jump process of a chemical system (which does not include reactions with delay) is a regular continuous-time HMC. In the

next section, we will see that this is no longer the case when the chemical system includes at least one reaction with delay. Then, how can we handle the problem? To answer this question, we recall the work done in the first Chapter of this thesis. In Section 2.3 we saw that a path representation of the reaction jump process can be given in terms of unit-rate Poisson processes.

Consider a chemical system of  $m$  chemical species  $S_1, \dots, S_m$  and a set of reactions  $\{\mathcal{R}_1, \dots, \mathcal{R}_n\}$  with propensities functions  $\{\alpha_1, \dots, \alpha_n\}$  and stoichiometric vectors  $\{v_1, \dots, v_n\}$ . Suppose that we know how the reaction jump process  $(X(s))$  is given in the time  $[0, t)$ . Then the dynamics of  $X(t)$  can be described as

$$(3.4) \quad X(t) = X(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(X(s)) ds \right) v[k],$$

where  $\mathcal{U}_k(\int_0^t \alpha_k(X(s)) ds)$  is the number of times that the reaction  $\mathcal{R}_k$  happens inside the time interval  $[0, t]$ , and the  $\mathcal{U}'_k$ s are independent unit rate Poisson process.

Now let  $\{t_i\}$  be the sequence of events of the counting process

$$\left\{ \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(X(s)) ds \right) \right\};$$

thus  $\forall \omega \in \Omega$ ,  $t_i(\omega)$  is the first time the function

$$t \rightarrow \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(X(s, \omega)) ds, \omega \right) \quad \text{is equal to } i.$$

Observe that we can interpret  $t_i$  as the time where the  $i$ th reaction occurs; also note that  $\forall i$  the chemical system can not produce reactions in the time interval  $(t_i, t_{i+1})$ . From  $\{t_i\} \cup \{\infty\}$  we can subtract a subsequence  $\{t_{k,j}\}_j$ , where  $\forall \omega \in \Omega$ , if  $t_{k,j}(\omega) \neq \infty$ , then  $t_{k,j}(\omega)$  is the first time the function

$$t \rightarrow \mathcal{U}_k \left( \int_0^t \alpha_k(X(s, \omega)) ds, \omega \right) \quad \text{is equal to } j.$$

that is, given the event  $\{t_{k,j} \neq \infty\}$ ,  $t_{k,j}$  is the time where the  $j$ th reaction of the type  $\mathcal{R}_k$  occurs and the chemical system can not produce reactions of the type  $\mathcal{R}_k$  in the time interval  $(t_{k,j}, t_{k,j+1})$ . It is important to note that the event

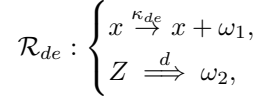
$$\{\alpha_k(X(s)) = 0 : \forall s \in [t_i, \infty)\}$$

may have probability greater than 0; then  $\forall \omega \in \{\alpha_k(X(s)) = 0 : \forall s \in [t_i, \infty)\}$ ,

$$t_{k,j}(\omega) = \infty \quad \forall j > \mathcal{U}_k \left( \int_0^{t_i(\omega)} \alpha_k(X(s, \omega)) ds, \omega \right).$$

Now suppose that this chemical system also induces a reaction with delay

$$\mathcal{R}_{de} = \mathcal{R}_{de}(d, \omega_1, \omega_2, \alpha_{de})$$



where  $\alpha_{de}$  is the propensity function of the chemical reaction part  $x \xrightarrow{\kappa_{de}} x + \omega_1$  of  $\mathcal{R}_{de}$ .

Note that a normal chemical reaction  $\mathcal{R}_k$  will change instantaneously the stoichiometry of the chemical system, but a reaction with delay  $\mathcal{R}_{de}$  will change the stoichiometry of the system in two times: the first one occurs instantaneously when the chemical reaction part  $x \rightarrow x + \omega_1$  of  $\mathcal{R}_{de}$  occurs, and the second one occurs  $d$  units of time after the first one happened when the complex  $Z$  decomposes. Then, we are interested in giving a description of the reaction jump process ( $\{Y(s)\}$ ) of this chemical system that can handle the changes in the stoichiometry of the system due to the decompositions of the complexes  $Z$ . Using (3.4) we can solve the problem due to stoichiometric changes that normal reactions induce. Since the first change that  $\mathcal{R}_{de}$  induces in the stoichiometry of the system is due to the normal reaction  $x \xrightarrow{\kappa_{de}} x + \omega_1$ , we can think that the dynamic of the chemical system at time  $t$  (given  $Y(s) \forall s \in [0, t)$ ) is

$$Y(t) = Y(0) + \left( \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(Y(s)) ds \right) v[k] \right) + \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \omega_1 + \beta(t)$$

where  $\mathcal{U}_{de}$  is a unit-rate Poisson process that is independent of  $\{\mathcal{U}_k : 1 \leq k \leq n\}$  and  $\beta(t)$  manage the stoichiometric changes (in the time interval  $[0, t]$ ) due to the decompositions of complexes  $Z$ . We can infer that  $\beta(t)$  will be a function of the number of complexes  $Z$  that have decomposed in the time interval  $[0, t]$  and the stoichiometric vector  $\omega_2$ . Every time a reaction  $x \xrightarrow{\kappa_{de}} x + \omega_1$  occurs a new complex  $Z$  is formed; therefore information about the counting process

$$\left\{ \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \right\},$$

can give us information about the times when the complexes  $Z$  will decompose. Let us be more precise; consider the sequence of events  $\{\hat{t}_{de,j}\}_j$  of the counting process

$$\left\{ \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \right\}.$$

Given the event  $\{\hat{t}_{de,1} < \infty\}$  (at least one delayed reaction  $\mathcal{R}_{de}$  occurs), then the first molecule of  $Z$  is formed at time  $\hat{t}_{de,1}$ , and at time  $\hat{t}_{de,1} + d$  this first complex will decompose ( $Z \xrightarrow{d} \omega_2$ ). It is clear then that if the event  $\{\hat{t}_{de,1} < \infty, \dots, \hat{t}_{de,j} < \infty\}$  occurs (at least  $j$  delayed reactions  $\mathcal{R}_{de}$  have occurred), for  $1 \leq i \leq j$ , the  $i$ th complex  $Z$  is formed at time  $\hat{t}_{de,i}$ , and at time  $\hat{t}_{de,i} + d$  this  $i$ th complex will decompose ( $Z \xrightarrow{d} \omega_2$ ). Therefore, the sequence of times at which the complexes

will decompose is  $\{\hat{t}_{de,j+d}\}_j$ , which has counting process given by

$$\left\{ \mathcal{U}_{de} \left( \int_0^{t-d} \alpha_{de}(Y(s)) ds \right) \right\},$$

because the following events are equal

$$\left\{ \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) = j \right\} = \{\hat{t}_{de,j} \leq t-d < \hat{t}_{de,j+1}\} = \{\hat{t}_{de,j+d} \leq t < \hat{t}_{de,j+1+d}\}.$$

Therefore, the reaction jump process of a chemical system with  $m$  chemical species,  $n$  chemical reactions and one reaction with delay can be written as

$$(3.5) \quad Y(t) = Y(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(Y(s)) ds \right) v[k] \\ + \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \omega_1 + \mathcal{U}_{de} \left( \int_0^{t-d} \alpha_{de}(Y(s)) ds \right) \omega_2$$

Let  $\star \in \{1, \dots, n, de\}$ . Since every unit-rate Poisson process  $\{\mathcal{U}_\star(u)\}_{u \geq 0}$  is of the form

$$\mathcal{U}_\star(u) = \sum_{j=0}^{\infty} j \chi_{[S_\star, j, S_\star, j+1)}(u),$$

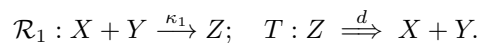
where  $\{S_\star, j\}_j$  is the sequence of events of  $\mathcal{U}_\star$  with  $S_\star, 0 = 0$ , then every possible solution of (3.5) has to be of the form

$$(3.6) \quad Y(t) = \sum_{j=0}^{\infty} Y_j \chi_{[\hat{t}_j, \hat{t}_{j+1})}(t)$$

where  $\hat{t}_0 = 0$  and for all  $j \in \mathbb{N}_0$   $Y_{j+1} = Y_j + w$  with  $w \in \{v_1, \dots, v_n, \omega_1, \omega_2\}$  and  $\{\hat{t}_j\}$  are the times where any stoichiometric change happens.

In chapter 3 we are going to introduce Anderson-Kurtz's algorithm for this type of chemical systems, which give us a way to simulate  $\{\hat{t}_j\}_j$  and  $\{Y_j\}$ , and we are going to compare this algorithm with Barrio's SSAD which is the standard tool used to simulate a chemical system with delays.

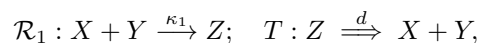
**Example 3.7.** (*Reaction Jump Process for the Binding and Unbinding Reaction with Delay*) Consider the reaction system with chemical species  $X, Y$ , and one excited complex  $Z$ , one reaction  $\mathcal{R}_1$ , and a physical transition  $T$  with delay  $d$



If the initial number of molecules of  $X$  and  $Y$  are  $x_0$  and  $y_0$  ( $x_0 \leq y_0$ ), then the reaction jump process is given by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \mathcal{U}_1 \left( \kappa_1 \int_0^t x(s)y(s) ds \right) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1 \left( \kappa_1 \int_0^{t-d} x(s)y(s) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

**3.2. A Semi-Markov Chemical Reaction System.** Consider the reaction system of example (3.3); that is one reaction  $\mathcal{R}_1$  and a physical transition  $T$  with delay  $d$





and initial state  $(1, y_0)^T$  with  $y_0 \in \mathbb{N}$ . Since at time 0 there are no molecules of complex  $Z$ , the first step of the chemical system is dictated by the chemical reaction  $\mathcal{R}_1$ ; then, at time  $\tau_1 \sim \exp(\kappa_1 y_0)$  the system moves instantaneously from  $(1, y_0)^T$  to  $(0, y_0 - 1)^T$ , and at the same time, the first excited complex  $Z$  is produced. Note that in the time interval  $[\tau_1, \tau_1 + d)$  the only molecule of  $X$  is binding with one molecule of  $Y$ , which implies that  $\forall s \in [\tau_1, \tau_1 + d)$   $(x(s), y(s))^T = (0, y_0 - 1)^T$ ; thus the propensity of reaction  $\mathcal{R}_1$  is equal to zero  $\forall s \in [\tau_1, \tau_1 + d)$ . A change in the propensity function of  $\mathcal{R}_1$  will occur at time  $\tau_1 + d$  when the first excited molecule  $Z$  decomposed in  $X$  and  $Y$ ; at this time the system moves instantaneously from  $(0, y_0 - 1)^T$  to  $(1, y_0)^T$  and the propensity of  $\mathcal{R}_1$  at time  $t_1 + d$  is equal to  $\kappa_1 y_0$ . At this point, there are no excited complex  $Z$  in the system, and the dynamics of the system is dictated again by  $\mathcal{R}_1$ .

Hence, if we let  $T_0 = 0, T_1 = T_0 + (\tau_1 + d), T_2 = T_1 + (\tau_2 + d), \dots, T_n = T_{n-1} + (\tau_n + d)$ , where  $\{\tau_k\}$  is a i.i.d. sequence with distribution  $\exp(\kappa_1 y_0)$ ; then we can write the reaction jump process as

$$(3.8) \quad (x(t), y(t))^T = \begin{cases} (1, y_0)^T & T_n \leq t < T_{n+1} - d \quad n \in \mathbb{N}_0, \\ (0, y_0 - 1)^T & T_{n+1} - d \leq t < T_{n+1} \quad n \in \mathbb{N}_0. \end{cases}$$

Note that  $\{T_n\}$  is the corresponding renewal process of  $\{\tau_n + d\}$ , where we can interpret  $T_n$  as the time where the  $n$ th excited complex  $Z$  decomposed or the time where the  $n$ th reaction with delay has finished. Also, note that (3.8) define a Semi-Markov process because the probability that the system moves from  $(0, y_0 - 1)^T$  to  $(1, y_0)^T$  follows a Dirac distribution centered on  $d$ .

In order to show that Anderson-Kurtz's equation for this chemical system

$$(3.9) \quad \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 1 \\ y_0 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^t \kappa_1(xy)(s) ds \right) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^{t-d} \kappa_1(xy)(s) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

is equivalent to (3.8), we must solve this equation. Since the unit-rate Poisson process

$$\{\mathcal{U}_1(u)\}_{u \geq 0}$$

is of the form

$$\mathcal{U}_1(u) = \sum_{j=0}^{\infty} j \chi_{[S_j, S_{j+1})}(u),$$

where  $\{S_j\}$  is the sequence of time events of  $\mathcal{U}_1$  and  $S_0 = 0$ , then every possible solution of (3.9) has to be of the form

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \sum_{j=0}^{\infty} \begin{pmatrix} x_j \\ y_j \end{pmatrix} \chi_{[\hat{t}_j, \hat{t}_{j+1})}(t)$$

where for  $j \geq 0$  the vector  $(x_{j+1}, y_{j+1})^T = (x_j, y_j)^T + v$  with

$$v \in \{(1, 1)^T, (-1, -1)^T\}$$

and  $\{\hat{t}_j\}_{j=0}^\infty$  are the times where any stoichiometric change happens; and we set  $\hat{t}_0=0$ .

Now we are going to solve explicitly (3.9) by constructing the sequences  $\{\hat{t}_j\}_{j=0}^\infty$  and  $\{(x_j, y_j)^T\}_{j \geq 0}$ . Note that  $(xy)(0) = y_0$ , because of the initial conditions, so that  $(xy)(s) = y_0 \forall s \in [0, \hat{t}_1)$ . Now then  $\hat{t}_1$  must satisfy by definition

$$\int_0^{\hat{t}_1} \kappa_1(xy)(s) ds = S_1,$$

and

$$\hat{t}_1 = \frac{S_1}{\kappa_1 y_0}$$

which implies that  $\hat{t}_1 \sim \exp(\kappa_1 y_0)$ . Observe that after substituting  $t = \hat{t}_1$  in (3.9) we obtain

$$\begin{pmatrix} x(\hat{t}_1) \\ y(\hat{t}_1) \end{pmatrix} = \begin{pmatrix} 1 \\ y_0 \end{pmatrix} + \mathcal{U}_1(S_1) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1(S_1 - d\kappa_1 y_0) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ y_0 - 1 \end{pmatrix}$$

so that  $(xy)(\hat{t}_1) = 0$ , the same identity  $(xy)(t) = 0$  happens for all  $t \geq \hat{t}_1$  until the integral

$$\int_0^{t-d} \kappa_1(xy)(s) ds$$

in the last term of (3.9) is equal to the event  $S_1$ . Since  $(xy)(s) = y_0 \forall s \in [0, \hat{t}_1)$ , then

$$\int_0^{t-d} \kappa_1(xy)(s) ds = \begin{cases} 0 & \text{if } t < d, \\ \kappa_1 y_0 (t-d) & \text{if } d \leq t \leq \hat{t}_1 + d. \end{cases}$$

and so (3.9) rewrites as follows  $\forall t \in [\hat{t}_1, \hat{t}_1 + d)$

$$\begin{aligned} (3.10) \quad \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \begin{pmatrix} 1 \\ y_0 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^t \kappa_1(xy)(s) ds \right) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^{t-d} \kappa_1(xy)(s) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ y_0 \end{pmatrix} + \mathcal{U}_1(S_1) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1(S_1 - d\kappa_1 y_0) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ y_0 - 1 \end{pmatrix}. \end{aligned}$$

This equation is consistent because

$$(xy)(s) = \begin{cases} y_0 & \text{if } s \in [0, \hat{t}_1), \\ 0 & \text{if } s \in [\hat{t}_1, \hat{t}_1 + d). \end{cases}$$

Moreover, substituting  $t = \hat{t}_1 + d$  into (3.9) and (3.10) yields

$$\begin{pmatrix} x(\hat{t}_1 + d) \\ y(\hat{t}_1 + d) \end{pmatrix} = \begin{pmatrix} 1 \\ y_0 \end{pmatrix} + \mathcal{U}_1(S_1) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1(S_1) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ y_0 \end{pmatrix},$$

because  $\mathcal{U}_1(S_1) = 1$ .

So that  $(xy)(\hat{t}_2) = y_0$  for the second stoichiometric time change  $\hat{t}_2 = \hat{t}_1 + d$ , the same identity  $(xy)(t) = y_0$  holds for all  $t \geq \hat{t}_2$  until the integral

$$\int_0^t \kappa_1(xy)(s) ds$$

is equal to the second time event  $S_2$ . If we continue with this process, then we can rewrite (3.9) as

$$(x(t), y(t))^T = \begin{cases} (1, y_0)^T & \hat{t}_{2n} \leq t < \hat{t}_{2n+1} \quad n \in \mathbb{N}_0, \\ (0, y_0 - 1)^T & \hat{t}_{2n+1} \leq t < \hat{t}_{2(n+1)} \quad n \in \mathbb{N}_0. \end{cases}$$

where  $\forall n \in \mathbb{N}_0$   $\hat{t}_{2n+1} - \hat{t}_{2n} \sim \exp(\kappa_1 y_0)$  independent of  $\hat{t}_1, \dots, \hat{t}_{2n}$  and

$$\hat{t}_{2(n+1)} = \hat{t}_{2n+1} + d.$$

Therefore, Anderson-Kurtz's representation (3.9) is equivalent to (3.8) after setting  $\hat{t}_{2n} = T_n$  and  $\hat{t}_{2n+1} = T_{n+1} - d$ .

We present and analyze now a second representation of the process (3.8). Let  $\{N([0, t])\}$  be the counting process for  $T_n$  (see (3.8)). Since  $T_n + d < T_{n+1}$  a.s., then (3.8) can be written a.s. as

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 1 \\ y_0 \end{pmatrix} + N([0, t+d]) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + N([0, t]) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

In particular,

$$x(t) = 1 - N([0, t+d]) + N([0, t])$$

and

$$\mathbb{P}(x(t) = 1) = 1 - E[N([0, t+d])] + E[N([0, t])]$$

because  $E[x(t)] = \mathbb{P}(x(t) = 1)$ . By Blackwell's theorem

$$\begin{aligned} & \lim_{t \rightarrow \infty} \mathbb{P}(x(t) = 1) \\ &= 1 - \lim_{t \rightarrow \infty} (E[N([0, t+d])] - E[N([0, t])]) \\ &= 1 - \frac{d}{E[\tau_1 + d]} \\ &= 1 - \frac{d}{\frac{1}{\kappa_1 y_0} + d} \\ &= 1 - \frac{\kappa_1 y_0 d}{1 + \kappa_1 y_0 d} \\ &= \frac{1}{1 + \kappa_1 y_0 d}. \end{aligned}$$

Then the steady state probabilities of finding the single molecule of  $X$  and the single molecule of complex  $Z$

$$p_\infty(X) := \lim_{t \rightarrow \infty} \mathbb{P}(x(t)=1),$$

$$p_\infty(Z) := \lim_{t \rightarrow \infty} \mathbb{P}(x(t)=0),$$

respectively, satisfies the following identity

$$(3.11) \quad p_\infty(X) = \frac{1}{1+\kappa_1 y_0 d}, \quad p_\infty(Z) = \frac{\kappa_1 y_0 d}{1+\kappa_1 y_0 d}.$$

Now that we have a clear characterization of the reaction jump process as a Semi-Markov process (see (3.8)), we can analyze the distribution of  $(x(t), y(t))^T$ ; from (3.8) we can write

$$\mathbb{P}((x(t), y(t))^T = (1, y_0)^T) = \sum_{k=0}^{\infty} \mathbb{P}(T_k \leq t < T_{k+1} - d).$$

which turns our question to calculate  $\mathbb{P}(T_k \leq t < T_{k+1} - d)$ . For  $k = 0$ , the result is clear, but what about  $k \geq 1$ ? First of all, suppose that all the random variables above has probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and remember that

$$T_{k+1} - d = T_k + \tau_{k+1},$$

where  $T_k$  and  $\tau_{k+1}$  are independent. For fixed  $t > 0$ , let  $D$  be the set

$$D = \{(x, y) \in \mathbb{R}^2 : 0 < y \leq t < x + y\}.$$

Then

$$\begin{aligned} & \mathbb{P}(T_k \leq t < T_k + \tau_{k+1}) \\ &= \int_{\Omega} \chi_D(\tau_{k+1}, T_k) dP \\ &= \int_{\mathbb{R}^2} \chi_D(x, y) d(P \circ (\tau_{k+1}, T_k)^{-1})(x, y) \\ &= \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \chi_D(x, y) d(P \circ \tau_{k+1}^{-1})(x) \right) d(P \circ T_k^{-1})(y) \\ &= \int_0^t \left( \int_{t-y}^{\infty} d(P \circ \tau_{k+1}^{-1})(x) \right) d(P \circ T_k^{-1})(y) \\ &= \int_0^t (\mathbb{P}(\tau_{k+1} > t-y)) d(P \circ T_k^{-1})(y) \\ &= \int_0^t (\exp(-\kappa_1 y_0(t-y))) d(P \circ T_k^{-1})(y). \end{aligned}$$

Therefore

$$\begin{aligned} & \mathbb{P}((x(t), y(t))^T = (1, y_0)^T) \\ &= \exp(-\kappa_1 y_0 t) + \sum_{k=1}^{\infty} \int_{(0, t]} \exp(-\kappa_1(t-y)) (d\mathbb{P} \circ T_k^{-1})(y). \end{aligned}$$

The last equality shows how complicated it can be to compute the distributions of the jump process, since we need for each  $k \in \mathbb{N}$  the  $k$ th convolution of the distribution function of  $\tau_1 + d$ .

## 4. SIMULATION OF THE REACTION JUMP PROCESS

In this chapter, we introduce Doob-Gillespie's Stochastic Simulation Algorithm (SSA), which is the basic tool for simulating the reaction jump process of a chemical system (without delayed reactions). As we are going to see, Doob-Gillespie's SSA is an implementation of the construction for the reaction jump process given in Section 2.2; therefore, we can give an interpretation to Doob-Gillespie's SSA in terms of a competition between reactions.

Thus, we consider the Anderson-Kurtz's equation for the chemical system  $\{\mathcal{R}_1, \dots, \mathcal{R}_n\}$

$$X(t) = X_0 + \sum_{j=1}^n \mathcal{U}_j \left( \int_0^t \alpha_j(X(s)) ds \right) v[j].$$

Following this equation and the solutions that generate, we obtain Anderson and Kurtz SSA. A natural question arises: Is Anderson and Kurtz SSA equivalent to Doob-Gillespie algorithm? The answer to this question is affirmative, and we introduce an elementary proof in this chapter.

**4.1. Doob-Gillespie Stochastic Simulation Algorithm (SSA).** Consider a reaction system of  $n$  reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n$  and  $m$  chemical species  $S_1, \dots, S_m$ . As we saw in Section 2.2, if the chemical system satisfies the fundamental hypothesis, then the dynamics of the system can be modeled as a continuous-time regular HMC  $\{X(t)\}_{t \geq 0}$  which satisfies

$$(4.1) \quad X(t) = \sum_{j=0}^{\infty} X_j \chi_{[t_j, t_{j+1})}(t),$$

where  $t_0=0$ , for all  $j \in \mathbb{N}_0$   $X_{j+1}=X_j+w$  with  $w \in \{v[1], \dots, v[n]\}$  and  $\{t_j\}_j$  are the times when a reaction occurs. These times satisfy

$$t_{j+1}-t_j \sim \exp \left( \sum_{k=1}^n \alpha_k(x_j) \right)$$

independent of  $t_0, \dots, t_j$  given that we know  $x_0, \dots, x_j$ .

Doob-Gillespie SSA is only an implementation of (4.1). For example, if  $X(0) = x_0$ , we simulate  $\tau_1 \sim \exp(\sum_{k=1}^n \alpha_k(x_0))$  and  $J_1$ , where

$$\mathbb{P}(J_1 = j_1 | X(0) = x_0) = \frac{\alpha_{j_1}(x_0)}{\sum_{k=1}^n \alpha_k(x_0)}$$

and  $J_1, \tau_1$  are independent given the event  $\{X(0) = x_0\}$ . We set  $t_1 = \tau_1$ , and if  $J_1=j_1, x_1=x_0+v[j_1]$ . At this point we generate  $\tau_2 \sim \exp(\sum_{k=1}^n \alpha_k(x_1))$  and  $J_2$  where

$$\mathbb{P}(J_2 = j_2 | X(0) = x_0, J_1 = j_1) = \frac{\alpha_{j_2}(x_1)}{\sum_{k=1}^n \alpha_k(x_1)}.$$

and  $\tau_2, J_2$  are  $\mathbb{P}(\cdot | X(0)=x_0, J_1=j_1)$ -independent. We set  $t_2=t_1+\tau_2$  and  $x_2=x_1+v_{j_2}$ . We continue with this process until the time reaches some time  $t_{\max}$ . The following lemmas will help us to simulate  $\tau$  and  $J$ .

**Lemma 4.2.** *If  $u \sim U(0, 1)$  and  $\lambda > 0$ , then  $-\frac{\ln(u)}{\lambda} \sim \exp(\lambda)$*

*Proof.* Let  $a > 0$ ; then

$$\begin{aligned} & \mathbb{P}\left(a < -\frac{\ln(u)}{\lambda}\right) \\ &= \mathbb{P}(-\lambda a > \ln(u)) \\ &= \mathbb{P}(\exp(-\lambda a) > u) \\ &= \exp(-\lambda a) \end{aligned}$$

□

**Lemma 4.3.** *Let  $\lambda_1, \dots, \lambda_k$  be positive real numbers and let  $u \sim U(0, 1)$ . Define*

$$I := \min \left\{ 1 \leq i \leq k : \frac{\sum_{j=1}^{i-1} \lambda_j}{\sum_{j=1}^k \lambda_j} \leq u < \frac{\sum_{j=1}^i \lambda_j}{\sum_{j=1}^k \lambda_j} \right\}.$$

*Then  $\mathbb{P}(I = i) = \frac{\lambda_i}{\sum_{j=1}^k \lambda_j}$ .*

The resulting algorithm can be summarized very simply in the following pseudocode, where an initial state  $X(0)$  is given:

**Algorithm 4.4.** *(Doob-Gillespie SSA)*

- (1) *Initialization.* Set  $x \leftarrow x_0$ ,  $t \leftarrow 0$
- (2) *Calculate  $\alpha_1(x), \dots, \alpha_n(x)$  and set*

$$\alpha \leftarrow (\alpha_1(x), \dots, \alpha_n(x)).$$

- (3) *Generate two independent random variables  $\tau \sim \exp(\sum_{k=1}^n \alpha_k)$  and*

$$u_1 \sim U(0, 1).$$

- (4) *Set  $j$  to be the smallest integer satisfying*

$$\sum_{k=1}^{j-1} \alpha_k \leq u_1 \sum_{k=1}^n \alpha_k < \sum_{k=1}^j \alpha_k.$$

- (5) *Update  $x \leftarrow x + v[j]$  and update  $t \leftarrow t + \tau$*
- (6) *Return to step 2.*

Since Doob-Gillespie SSA is a clear consequence of (4.1), and (4.1) is a progressively measurable process, we can conclude that Doob-Gillespie SSA preserve the measurability of the reaction jump process.

Now we are interested in giving an interpretation of Algorithm 4.4 as a competition between the chemical reactions in the chemical system. Suppose that at time  $t$  the system satisfies  $X(t) = x$ . Following the construction given in Section 2.2 we need

to consider the set of independent HPP

$$\{N_{x,x+v[k]} : \forall 1 \leq k \leq n \quad N_{x,x+v[k]} \text{ has intensity } \alpha_k(x)\}.$$

If for each  $1 \leq k \leq n$ , we set  $T_{k,1}^x$  as the first event time of  $N_{x,x+v[k]}$ , then the event “the next reaction is  $\mathcal{R}_j$ ” can be written as

$$\{T_{j,1}^x < T_{l,1}^x : l \neq j\}.$$

Since for each  $1 \leq k \leq n$ ,  $T_{k,1}^x \sim \exp(\alpha_k(x))$ , and the  $\{T_{k,1}^x\}_k$  are independent, the event  $\{T_{j,1}^x < T_{l,1}^x : l \neq j\}$  is equal to

$$\{T_{j,1}^x = \min\{T_{l,1}^x : \forall 1 \leq l \leq n\}\},$$

except for a set of zero probability. Let

$$(4.5) \quad \tau = \min\{T_{k,1}^x : \forall 1 \leq k \leq n\}$$

and observe that we can interpret  $\tau$  as the time until the next reaction, given that the system is in state  $x$ . If  $J$  is the index where the minimum is reached, then  $J$  is the type of reaction that occurs.

Now we are going to prove that  $\tau$  and  $J$  have the same distributions as in Algorithm 4.4. We begin with a result known as The freezing lemma (see [2]).

**Theorem 4.6.** *(The freezing lemma) Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and  $\mathcal{G}$  and  $\mathcal{D}$  independent sub- $\sigma$ -algebras of  $\mathcal{F}$ . Let  $X$  be a  $\mathcal{D}$ -measurable random variable taking values in the measurable space  $(E, \mathcal{E})$  and*

$$\Psi : E \times \Omega \rightarrow \mathbb{R}$$

*an  $\mathcal{E} \times \mathcal{G}$ -measurable function such that  $\omega \mapsto \Psi(X(\omega), \omega)$  is integrable. Then*

$$E[\Psi(X, \cdot) | \mathcal{D}] = \Theta(X),$$

*where  $\Theta(x) = E[\Psi(x, \cdot)]$ .*

All the results in this thesis can be demonstrated in an elementary way but we want to show how The freezing lemma can be used.

**Lemma 4.7.** *Let  $i_1, \dots, i_k$  be in  $\mathbb{N}$  and let  $X_{i_1}, \dots, X_{i_k}$  be  $k$  independent random variables with distribution  $\exp(\lambda_{i_1}), \dots, \exp(\lambda_{i_k})$ , respectively. If  $U = \min\{X_{i_1}, \dots, X_{i_k}\}$ , then*

$$U \sim \exp(\lambda_{i_1} + \dots + \lambda_{i_k}).$$

*Proof.* Take  $a \in [0, \infty)$

$$\begin{aligned} \mathbb{P}(U > a) &= \mathbb{P}(X_{i_1} > a, \dots, X_{i_k} > a) \\ &= \mathbb{P}(X_{i_1} > a) \cdots \mathbb{P}(X_{i_k} > a) \\ &= \exp(-\lambda_{i_1} a) \cdots \exp(-\lambda_{i_k} a) \\ &= \exp(-(\lambda_{i_1} + \dots + \lambda_{i_k}) a). \end{aligned}$$

□

**Lemma 4.8.** *Let  $X_1, \dots, X_n$  be  $n$  independent random variables with distribution  $\exp(\lambda_1), \dots, \exp(\lambda_n)$ , respectively, and well adapted to the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $Z = \min\{X_1, \dots, X_n\}$  and let  $J$  be such that  $Z = X_J$ . Then  $J$  and  $Z$  are two independent random variables and satisfy:  $Z \sim \exp(\lambda_1 + \dots + \lambda_n)$  and*

$$\mathbb{P}(J = i, Z > a) = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n} \exp(-(\lambda_1 + \dots + \lambda_n)a).$$

*Proof.* Let  $U = \min\{X_j : j \neq i\}$ ; then  $\mathbb{P}(J = i, Z > a) = \mathbb{P}(a < X_i \leq U)$ . Now let  $\Psi : (\mathbb{R}) \times \Omega \rightarrow \mathbb{R}$  be given by

$$\Psi(x, \omega) := \begin{cases} 1 & a < x \leq U(\omega), \\ 0 & \text{i.o.c.} \end{cases}$$

Note that if  $x \in \mathbb{R}$  is fixed and  $a < x$ , then

$$\{\omega \in \Omega : \Psi(x, \omega) = 1\} = \{x \leq U\};$$

and if  $x \leq a$ ,  $\{\omega \in \Omega : \Psi(x, \omega) = 0\} = \Omega$ . Thus, in every case the function  $\omega \rightarrow \Psi(x, \omega)$  is  $\sigma(U)$ -measurable. Also note that for each  $\omega$ ,  $x \rightarrow \Psi(x, \omega)$  is left continuous. Then  $\Psi$  is  $\mathcal{B}(\mathbb{R}) \times \sigma(U)$ -measurable. Using the freezing lemma (Theorem 4.6), the random variable  $\omega \rightarrow \Psi(X_i(\omega), \omega)$  satisfies

$$E[\Psi(X_i, \cdot) | X_i] = \Theta(X_i),$$

where  $\Theta(x) = E[\Psi(x, \cdot)]$ , that is

$$\Theta(x) = \begin{cases} \mathbb{P}(x \leq U) & a < x, \\ 0 & x \leq a. \end{cases}$$

Note

$$\{a < X_i \leq U\} = \{\omega \in \Omega : \Psi(X_i(\omega), \omega) = 1\}.$$

Therefore

$$\begin{aligned} & \mathbb{P}(a < X_i \leq U) \\ &= \mathbb{P}(\Psi(X_i, \cdot)) \\ &= E[\Psi(X_i, \cdot)] \\ &= E[E[\Psi(X_i, \cdot) | X_i]] \\ &= E[\Theta(X_i)] \\ &= \int_{\mathbb{R}} \Theta(x) d(P \circ X_i^{-1}(x)) \\ &= \int_a^\infty \mathbb{P}(x \leq U) \lambda_i \exp(-\lambda_i x) dx \\ &= \int_a^\infty \exp(-(\sum_{j \neq i} \lambda_j)x) \lambda_i \exp(-\lambda_i x) dx \\ &= \int_a^\infty \exp(-(\sum_j \lambda_j)x) \lambda_i dx \\ &= \frac{\lambda_i}{\sum_j \lambda_j} \exp(-(\sum_j \lambda_j)a). \end{aligned}$$



From Lemma 4.7  $\mathbb{P}(x \leq U) = \exp(-(\sum_{j \neq i} \lambda_j)x)$ . Hence

$$(4.9) \quad \mathbb{P}(J = i, Z > a) = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n} \exp(-(\lambda_1 + \dots + \lambda_n)a),$$

and by continuity of the measure

$$\begin{aligned} & \mathbb{P}(J = i) \\ &= \lim_{a \downarrow 0} \mathbb{P}(J = i, Z > a) \\ &= \lim_{a \downarrow 0} \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n} \exp(-(\lambda_1 + \dots + \lambda_n)a) \\ &= \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}. \end{aligned}$$

Therefore,  $\forall 1 \leq i \leq n$

$$(4.10) \quad \mathbb{P}(J = i) = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}.$$

Since Lemma 4.7 implies that  $Z \sim \exp(\lambda_1 + \dots + \lambda_n)$ , then equations (4.9) and (4.10) imply that  $J$  and  $Z$  are independent.  $\square$

It follows from lemma (4.8) that  $\tau$  and  $J$  as in (4.5) have the same distributions as in Doob-Gillespie SSA (Algorithm 4.4). Also we can write Doob-Gillespie SSA in the following way.

**Algorithm 4.11.** (*Doob-Gillespie SSA Competitive version*)

- (1) *Initialization.* Set  $x \leftarrow x_0$ ,  $t \leftarrow 0$
- (2) *Calculate*  $\alpha_1(x), \dots, \alpha_n(x)$  *and set*

$$\alpha \leftarrow (\alpha_1(x), \dots, \alpha_n(x)).$$

- (3) *Generate*  $n$  *independent random variables*  $\tau_1, \dots, \tau_n$  *with*  $\tau_k \sim \exp(\alpha_k)$ .
- (4) *Set*  $\tau = \min\{\tau_1, \dots, \tau_n\}$ .
- (5) *Set*  $K = \{1 \leq k \leq n : \tau_k = \tau\}$
- (6) *Update*  $x$  *as*

$$x \leftarrow x + \sum_{k \in K} v[k]$$

- (7) *Update*  $t$  *as*  $t \leftarrow t + \tau$
- (8) *Return to step 2.*

Note that  $K$  in (5) has cardinality  $|K| > 1$  with probability zero.

Although Algorithm 4.11 has a clear interpretation as a competition between reactions, it is not as efficient as Algorithm 4.4 because every step of 4.11 needs to generate  $n$  exponential and independent random variables, while 4.4 only generates two independent random variables in each step.

Consider again the construction given in Section 2.2. Suppose that  $X(0) = x_0$ ; at this point we model the competition between the reactions through a competition of independent Poisson processes

$$\{N_{x_0, x_0+v[k]} : \forall 1 \leq k \leq n \quad N_{x_0, x_0+v[k]} \text{ has intensity } \alpha_k(x_0)\}.$$

From Lemma 4.8 we know that the index of the first reaction  $J_1$  and the time until the first reaction  $\tau_1$  are independent, and we know their distributions.

Suppose that the event  $\{X(0) = x_0, J_1 = j_1\}$  has positive probability. Set  $x_1 = x_0 + v[j_1]$  and note that

$$\{X(0) = x_0, J_1 = j_1\} = \{X(0) = x_0, X_1 = x_1\}.$$

Given the event  $\{X(0) = x_0, J_1 = j_1\}$  we consider a set of independent Poisson processes

$$\{N_{x_1, x_1+v[k]} : \forall 1 \leq k \leq n \quad N_{x_1, x_1+v[k]} \text{ has intensity } \alpha_k(x_1)\}$$

and independent of

$$\{N_{x_0, x_0+v[k]} : \forall 1 \leq k \leq n \quad N_{x_0, x_0+v[k]} \text{ has intensity } \alpha_k(x_0)\}.$$

If  $\tau_1 = \min\{T_{k,1}^{x_0} : 1 \leq k \leq n\}$ ,  $\tau_2 = \min\{T_{k,1}^{x_1} : 1 \leq k \leq n\}$  (see 4.5) and  $J_2$  is as in Lemma (4.8), then

$$\tau_1, \tau_2, J_2 \text{ are } \mathbb{P}(\cdot | X(0) = x_0, J_1 = j_1) \text{ - independent.}$$

Therefore, if we continue with this process, then given the event

$$\{X(0) = x_0, J_1 = j_1, \dots, J_q = j_q\} = \{X(0) = x_0, X_1 = x_1, \dots, X_q = x_q\}$$

the following sets of Poisson process are independent

$$\{\{N_{x_i, x_i+v[k]} : \forall 1 \leq k \leq n \quad N_{x_i, x_i+v[k]} \text{ has intensity } \alpha_k(x_i)\} : 0 \leq i \leq q\}.$$

This implies

**Theorem 4.12.** *Let  $q \in \mathbb{N}$ . Given the chemical reaction's chain until time  $q$*

$$\{X(0), J_1 = j_1, \dots, J_q = j_q\},$$

*the times between chemical reactions and the index of the next reaction are independent, that is*

$$\tau_1, \dots, \tau_q, J_{q+1}$$

*are  $\mathbb{P}(\cdot | X(0), J_1 = j_1, \dots, J_q = j_q)$ -independent*

With this in mind we can rewrite Doob-Gillespie SSA in the following way

**Algorithm 4.13.** *(Doob-Gillespie SSA Chemical Reaction's Chain version)*

(1) *Initialization. Set  $x \leftarrow x_0$ ,  $t \leftarrow 0$ ,  $step \leftarrow 0$ ,  $Nsteps \leftarrow q$  and*

$$Statemat \leftarrow 0_{m \times Nsteps+1}.$$

(2) *Update Statemat as  $Statemat(:, 1) \leftarrow x$*

(3) *Calculate  $\alpha_1(x), \dots, \alpha_n(x)$  and set*

$$\alpha \leftarrow (\alpha_1(x), \dots, \alpha_n(x)).$$

(4) Generate  $u \sim U(0, 1)$ . Set  $j$  to be the smallest integer satisfying

$$\sum_{k=1}^{j-1} \alpha_k \leq u \sum_k \alpha_k < \sum_{k=1}^j \alpha_k.$$

(5) Update  $x \leftarrow x + v[j]$ .

(6) Update  $step \leftarrow step + 1$ .

(7) Update *Statemat* as  $Statemat(:, step + 1) \leftarrow x$ .

(8) While  $step \leq Nsteps$  Return to step (3). Otherwise go to step (9).

(9) Generate  $Nsteps + 1$  independent random variables  $\tau_1, \dots, \tau_{Nsteps + 1}$  such that

$$\tau_i \sim \exp\left(\sum_{k=1}^n \alpha_k(Statemat(:, i))\right).$$

**4.2. Anderson-Kurtz's SSA.** Now we put our attention back to Anderson-Kurtz's equation for the chemical system given in Section 2.3.

$$(4.14) \quad X(t) = X(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(X(s)) ds \right) v[k].$$

As we said in Section 2.3, every solution of (4.14) has to be of the form

$$(4.15) \quad X(t) = \sum_{j=0}^{\infty} X_j \chi_{[t_j, t_{j+1})}(t),$$

where  $t_0 = 0$ , for all  $j \in \mathbb{N}_0$  the vector  $X_{j+1} = X_j + w$  with  $w \in \{v[1], \dots, v[n]\}$ , and  $\{t_j\}_j$  are the times when a reaction occurs.

Following (4.14) and (4.15), we get Anderson-Kurtz's SSA which allows us to calculate and simulate the sequences  $\{X_j\}_j, \{t_j\}_j$  in (4.15). Similar to Doob-Gillespie's SSA, Anderson-Kurtz's SSA give us another way to simulate the times between reactions ( $\{\tau_j = t_{j+1} - t_j\}_j$ ) and the type of reaction that occurs ( $\{J_j\}_j$ ); see Algorithm 4.21. Our objective is to show that Anderson-Kurtz's SSA and Doob-Gillespie SSA are equivalent. To do so, we need the following results:

(1) For every  $k \in \mathbb{N}$ , if we know the first  $k$  reactions, then

$$\tau_1, \dots, \tau_k, \tau_{k+1}, J_{k+1}$$

are independent. More precisely, given the event

$$\{X(0) = x_0, J_1 = j_1, \dots, J_k = j_k\},$$

$\tau_1, \dots, \tau_k, \tau_{k+1}, J_{k+1}$  are independent, for each  $1 \leq i \leq k + 1$

$$\tau_i \sim \exp\left(\sum_{k=1}^n \alpha_k \left(x_0 + \sum_{l=1}^{i-1} v[j_l]\right)\right)$$

and

$$\mathbb{P}(J_{k+1} = j_{k+1} | X(0) = x_0, J_1 = j_1, \dots, J_k = j_k) = \frac{\alpha_{j_{k+1}}(x_0 + \sum_{l=1}^k v[j_l])}{\sum_{k=1}^n \alpha_k(x_0 + \sum_{l=1}^k v[j_l])}.$$

Note that (1) implies that Anderson-Kurtz's equation (4.14) has a unique solution and this solution is given by (4.1).

Now we introduce the machinery that will help us to prove (1).

**Proposition 4.16.** *Let  $X_1, \dots, X_n$  be i.i.d. random variables with distribution  $\exp(1)$ , well adapted to the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $\alpha_1, \dots, \alpha_n$  be nonnegative numbers; let*

$$Z = \min \left\{ \frac{X_1}{\alpha_1}, \dots, \frac{X_n}{\alpha_n} \right\},$$

and  $J$  such that  $Z = \frac{X_J}{\alpha_J}$  (if for some  $j$   $\alpha_j = 0$ , then  $\frac{X_j}{\alpha_j} \sim \exp(0)$ ; that is, it is equal to  $\infty$  with probability 1).

Fix  $i$  such that  $\mathbb{P}(J = i) > 0$ , consider  $k \in \{1, \dots, n-1\}$  index  $i_1, \dots, i_k$  different from  $i$ . Then, for any  $a, a_1, \dots, a_k$  nonnegative real numbers

$$\mathbb{P}(X_{i_1} - \alpha_{i_1} Z > a_1, \dots, X_{i_k} - \alpha_{i_k} Z > a_k | J = i, Z > a) = \exp(a_1) \dots \exp(a_k).$$

Therefore given the event  $\{J = i, Z > a\}$ , the set  $\{X_k - \alpha_k Z : k \neq i\}$  is a i.i.d. sequence with distribution  $\exp(1)$ .

*Proof.* Notice that  $a_l \neq 0$ .

$$\begin{aligned} & \mathbb{P}(X_{i_1} - \alpha_{i_1} Z > a_1, \dots, X_{i_k} - \alpha_{i_k} Z > a_k, J = i, Z > a) \\ &= \mathbb{P}(X_{i_1} - \alpha_{i_1} \frac{X_i}{\alpha_i} > a_1, \dots, X_{i_k} - \alpha_{i_k} \frac{X_i}{\alpha_i} > a_k, J = i, \frac{X_i}{\alpha_i} > a) \\ &= \mathbb{P}(X_{i_1} - \alpha_{i_1} \frac{X_i}{\alpha_i} > a_1, \dots, X_{i_k} - \alpha_{i_k} \frac{X_i}{\alpha_i} > a_k, \frac{X_j}{\alpha_j} \geq \frac{X_i}{\alpha_i} \quad \forall j \neq i, \frac{X_i}{\alpha_i} > a) \\ &= \mathbb{P}(X_{i_1} > a_1 + \frac{\alpha_{i_1}}{\alpha_i} X_i, \dots, X_{i_k} > a_k + \frac{\alpha_{i_k}}{\alpha_i} X_i, \quad X_j \geq \frac{\alpha_j}{\alpha_i} X_i \quad \forall j \notin \{i_1, \dots, i_k, i\}, X_i > \alpha_i a) \end{aligned}$$

Let  $\Psi : \mathbb{R} \times \Omega \rightarrow \mathbb{R}$  be given by

$$\Psi(x, \omega) = \begin{cases} 1 & \text{if } X_{i_1}(\omega) > a_1 + \frac{\alpha_{i_1}}{\alpha_i} x, \dots, X_{i_k}(\omega) > a_k + \frac{\alpha_{i_k}}{\alpha_i} x, X_j(\omega) \geq \frac{\alpha_j}{\alpha_i} x \quad \forall j \notin \{i_1, \dots, i_k, i\}, x > \alpha_i a \\ 0 & \text{i.o.c.} \end{cases}$$

Fixed  $x \in \mathbb{R}$  it is clear that

$$\Psi(x, \cdot)^{-1}(1) = \begin{cases} \{X_{i_1} > a_1 + \frac{\alpha_{i_1}}{\alpha_i} x, \dots, X_{i_k} > a_k + \frac{\alpha_{i_k}}{\alpha_i} x, X_j \geq \frac{\alpha_j}{\alpha_i} x \quad \forall j \notin \{i_1, \dots, i_k, i\}\} & \text{if } x > \alpha_i a \\ \emptyset & \text{if } x \leq \alpha_i a. \end{cases}$$

Therefore, for any  $x \in \mathbb{R}$ ,  $\sigma(\Psi(x, \cdot)) \subset \sigma(X_j : j \neq i)$ ; i.e. for any  $x \in \mathbb{R}$ ,  $\sigma(\Psi(x, \cdot))$  is  $\sigma(X_j : j \neq i)$ -measurable. Now, for fixed  $\omega$   $x \rightarrow \Psi(x, \omega)$  is right continuous. Then  $\Psi$  is  $\mathcal{B}(\mathbb{R}) \times \sigma(X_j : j \neq i)$ -measurable. By the freezing lemma

$$E[\Psi(X_i, \cdot) | X_i] = \Theta(X_i)$$

where  $\forall x \in \mathbb{R}$ ,

$$\begin{aligned}
& \Theta(x) \\
&= E[\Psi(x, \cdot)] \\
&= \begin{cases} \mathbb{P}(X_{i_1} > a_1 + \frac{\alpha_{i_1}}{\alpha_i} x, \dots, X_{i_k} > a_k + \frac{\alpha_{i_k}}{\alpha_i} x, X_j \geq \frac{\alpha_j}{\alpha_i} x \quad \forall j \notin \{i_1, \dots, i_k, i\}) & \text{if } x > \alpha_i a \\ 0 & \text{if } x \leq \alpha_i a \end{cases} \\
&= \begin{cases} \mathbb{P}(X_{i_1} > a_1 + \frac{\alpha_{i_1}}{\alpha_i} x) \cdots \mathbb{P}(X_{i_k} > a_k + \frac{\alpha_{i_k}}{\alpha_i} x) \prod_{j \notin \{i_1, \dots, i_k, i\}} \mathbb{P}(X_j \geq \frac{\alpha_j}{\alpha_i} x) & \text{if } x > \alpha_i a \\ 0 & \text{if } x \leq \alpha_i a \end{cases} \\
&= \begin{cases} \exp(-(a_1 + \frac{\alpha_{i_1}}{\alpha_i} x)) \cdots \exp(-(a_k + \frac{\alpha_{i_k}}{\alpha_i} x)) \prod_{j \notin \{i_1, \dots, i_k, i\}} \exp(-(\frac{\alpha_j}{\alpha_i} x)) & \text{if } x > \alpha_i a \\ 0 & \text{if } x \leq \alpha_i a \end{cases} \\
&= \begin{cases} \exp(-a_1) \cdots \exp(-a_k) \exp(-(\frac{\sum_{j \neq i} \alpha_j}{\alpha_i} x)) & \text{if } x > \alpha_i a \\ 0 & \text{if } x \leq \alpha_i a. \end{cases}
\end{aligned}$$

Since the expectation of the random variable  $\omega \rightarrow \Psi(X_i(\omega), \omega)$  satisfies

$$E[\Psi(X_i, \cdot)] = \mathbb{P}(X_{i_1} > a_1 + \frac{\alpha_{i_1}}{\alpha_i} X_i, \dots, X_{i_k} > a_k + \frac{\alpha_{i_k}}{\alpha_i} X_i, \quad X_j \geq \frac{\alpha_j}{\alpha_i} X_i \quad \forall j \notin \{i_1, \dots, i_k, i\}, X_i > \alpha_i a),$$

then

$$\begin{aligned}
& \mathbb{P}(X_{i_1} > a_1 + \frac{\alpha_{i_1}}{\alpha_i} X_i, \dots, X_{i_k} > a_k + \frac{\alpha_{i_k}}{\alpha_i} X_i, \quad X_j \geq \frac{\alpha_j}{\alpha_i} X_i \quad \forall j \notin \{i_1, \dots, i_k, i\}, X_i > \alpha_i a) \\
&= E[E[\Psi(X_i, \cdot) | X_i]] \\
&= E[\Theta(X_i)] \\
&= \int_0^\infty \Theta(x) d\mathbb{P} \circ X_i^{-1}(x) \\
&= \exp(-a_1) \cdots \exp(-a_k) \int_{\alpha_i a}^\infty \exp(-(\frac{\sum_{j \neq i} \alpha_j}{\alpha_i} x)) \exp(-x) dx \\
&= \exp(-a_1) \cdots \exp(-a_k) \int_{\alpha_i a}^\infty \exp(-(\frac{\sum_j \alpha_j}{\alpha_i} x)) dx.
\end{aligned}$$

Now

$$\begin{aligned}
& \mathbb{P}(X_{i_1} - \alpha_{i_1} Z > a_1, \dots, X_{i_k} - \alpha_{i_k} Z > a_k | J = i, Z > a) \\
&= \frac{1}{\mathbb{P}(J=i, Z>a)} \mathbb{P}(X_{i_1} > a_1 + \frac{\alpha_{i_1}}{\alpha_i} X_i, \dots, X_{i_k} > a_k + \frac{\alpha_{i_k}}{\alpha_i} X_i, \quad X_j \geq \frac{\alpha_j}{\alpha_i} X_i \quad \forall j \notin \{i_1, \dots, i_k, i\}, X_i > \alpha_i a) \\
&= \frac{1}{\mathbb{P}(J=i) \mathbb{P}(Z>a)} [\exp(-a_1) \cdots \exp(-a_k)] \int_{\alpha_i a}^\infty \exp(-(\frac{\sum_j \alpha_j}{\alpha_i} x)) dx \\
&= \exp((\sum_j \alpha_j) a) (\frac{\sum_j \alpha_j}{\alpha_i}) [\exp(-a_1) \cdots \exp(-a_k)] \int_{\alpha_i a}^\infty \exp(-(\frac{\sum_j \alpha_j}{\alpha_i} x)) dx \\
&= [\exp(-a_1) \cdots \exp(-a_k)] \exp((\sum_j \alpha_j) a) \left[ \int_{\alpha_i a}^\infty \frac{\sum_j \alpha_j}{\alpha_i} \exp(-(\frac{\sum_j \alpha_j}{\alpha_i} x)) dx \right] \\
&= [\exp(-a_1) \cdots \exp(-a_k)] \exp((\sum_j \alpha_j) a) \exp(-(\frac{\sum_j \alpha_j}{\alpha_i}) \alpha_i a) \\
&= \exp(-a_1) \cdots \exp(-a_k).
\end{aligned}$$

From Lemma 4.8,  $J$  and  $Z$  are independent and  $Z \sim \exp(\sum_j \alpha_j)$  □

**Corollary 4.17.** *Let  $X_1, \dots, X_n$  be i.i.d. random variables with distribution  $\exp(1)$ , well adapted to the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $\alpha_1, \dots, \alpha_n$  be nonnegative numbers; let*

$$Z = \min \left\{ \frac{X_1}{\alpha_1}, \dots, \frac{X_n}{\alpha_n} \right\},$$

and  $J$  such that  $Z = \frac{X_J}{\alpha_J}$  (if for some  $j$   $\alpha_j = 0$ , then  $\frac{X_j}{\alpha_j} \sim \exp(0)$ ; that is, it is equal to  $\infty$  with probability 1).

Fix  $i$  such that  $\mathbb{P}(J = i) > 0$ , consider  $k \in \{1, \dots, n-1\}$  index  $i_1, \dots, i_k$  different from  $i$ . Then, for  $a_1, \dots, a_k$  positive numbers

$$\mathbb{P}(X_{i_1} - \alpha_{i_1} Z > a_1, \dots, X_{i_k} - \alpha_{i_k} Z > a_k | J = i) = \exp(a_1) \dots \exp(a_k).$$

Therefore, given the event  $\{J = i\}$ ,  $\{X_k - \alpha_k Z : 1 \leq k \leq n\}$  is a i.i.d sequence with distribution  $\exp(1)$  except for  $k = i$ .

*Proof.* Since  $Z$  has exponential distribution, by continuity of the measure

$$\lim_{a \downarrow 0} \mathbb{P}(\cdot | J = i, Z > a) = \mathbb{P}(\cdot | J = i)$$

Then by proposition (4.16),

$$\mathbb{P}(X_{i_1} - \alpha_{i_1} Z > a_1, \dots, X_{i_k} - \alpha_{i_k} Z > a_k | J = i) = \exp(a_1) \dots \exp(a_k).$$

□

**Corollary 4.18.** Let  $X_1, \dots, X_n$  be i.i.d. random variables with distribution  $\exp(1)$ , well adapted to the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $\alpha_1, \dots, \alpha_n$  be nonnegative numbers; let

$$Z = \min \left\{ \frac{X_1}{\alpha_1}, \dots, \frac{X_n}{\alpha_n} \right\},$$

and  $J$  such that  $Z = \frac{X_J}{\alpha_J}$  (if for some  $j$   $\alpha_j = 0$ , then  $\frac{X_j}{\alpha_j} \sim \exp(0)$ ; that is, it is equal to  $\infty$  with probability 1).

Fix  $i$  such that  $\mathbb{P}(J = i) > 0$ . Then given the event  $\{J = i\}$ , the set  $\{X_k - \alpha_k Z : 1 \leq k \leq n\}$  is independent of  $Z$ .

*Proof.* Let  $a > 0$ ,  $b > 0$  and  $k \neq i$ . By Corollary 4.17 we have

$$\begin{aligned} & \mathbb{P}(X_k - \alpha_k Z > a, Z > b | J = i) \\ &= \frac{1}{\mathbb{P}(J=i)} \mathbb{P}(X_k - \alpha_k Z > a, Z > b, J = i) \\ &= \frac{\mathbb{P}(Z > b, J=i)}{\mathbb{P}(J=i)} \mathbb{P}(X_k - \alpha_k Z > a | Z > b, J = i) \\ &= \mathbb{P}(X_k - \alpha_k Z > a | J = i) \mathbb{P}(Z > b | J = i) \end{aligned}$$

□

Consider again a reaction system of  $n$  reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n$  and  $m$  chemical species  $S_1, \dots, S_m$ . As we saw in Section 2.3, given the dynamics of the reaction jump process in the time interval  $[0, t)$ , the dynamic of the system at time  $t$  is given by Anderson-Kurtz's equation

$$(4.19) \quad X(t) = X(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(X(s)) ds \right) v[k].$$

which has solutions of the form

$$(4.20) \quad X(t) = \sum_{j=0}^{\infty} x_j \chi_{[t_j, t_{j+1})}(t)$$

where  $t_0=0$ , for all  $j \in \mathbb{N}_0$  the vector  $x_{j+1}=x_j+w$  for some

$$w \in \{v[1], \dots, v[n]\}.$$

The set  $\{t_j\}$  are the times when a reaction occurs.

Consider the set of independent unit-rate Poisson processes  $\mathcal{U}_1 \dots \mathcal{U}_n$  and suppose that these processes are independent of the initial condition  $X(0)$ ; suppose that each Poisson process  $\mathcal{U}_k$  is well adapted to the probability  $(\Omega, \mathcal{F}, \mathbb{P})$ ; that is, for each  $1 \leq k \leq n$

$$\sigma(\mathcal{U}_k(t) : 0 < t) \subset \mathcal{F}$$

For each  $1 \leq k \leq n$ , let  $\{S_{k,j}\}_{j \geq 0}$  be the sequence of events of the unit-rate Poisson process  $\mathcal{U}_k$ . Then, for each  $j \in \mathbb{N}$

$$S_{k,j} = \inf\{t > 0 : \mathcal{U}_k(t) = j\}, \quad S_{k,0}=0.$$

For each  $l \in \mathbb{N}$  the sequence of random variables

$$S_{k,1}-S_{k,0}; S_{k,2}-S_{k,1}; \dots; S_{k,l}-S_{k,l-1}$$

are i.i.d. with distribution  $\exp(1)$ .

Suppose that the chemical system satisfies  $X(0) = x_0$  with positive probability and define a probability measure  $\mathbb{P}_0$  by

$$\mathbb{P}_0(\cdot) = \mathbb{P}(\cdot | X(0) = x_0).$$

We are interested in the time  $t_1$  when the system does a first transition to another state due to some reaction  $\mathcal{R}_1, \dots, \mathcal{R}_n$ . Therefore, given the  $\{X(0) = x_0\}$  and the fact that each solution of (4.19) has the form (4.20), then for each  $1 \leq k \leq n$  we must have the following equality

$$\int_0^{t_1} \alpha_k(X(s)) ds = t_1 \alpha_k(x_0).$$

Also, from (4.19) note that the first reaction will happen when some  $t_1 \alpha_k(x_0)$  hits the first event time of  $\mathcal{U}_k$ , but due to the stoichiometry changes for the first reaction  $t_1$  must satisfies

$$t_1 = \min_{1 \leq k \leq n} \left\{ \frac{S_{k, l_k(1)} - \alpha_k(x_{-1}) \tau_0}{\alpha_k(x_0)} \right\}$$

where  $l_k(1)$  is the first event of  $\mathcal{U}_k$  such that  $S_{k, l_k(1)} - \alpha_k(x_{-1}) \tau_0 > 0$ , and we define  $x_{-1} := 0$  and  $\tau_0 := 0$ . Then for each  $1 \leq k \leq n$

$$\alpha_k(x_{-1}) \tau_0 = 0.$$

Thus, it is clear that a.s.  $\forall 1 \leq k \leq n, l_k(1) = 1$  and

$$t_1 = \min_{1 \leq k \leq n} \left\{ \frac{S_{k,1}}{\alpha_k(x_0)} \right\}.$$

Since the Poisson process  $\mathcal{U}_1, \dots, \mathcal{U}_k$  are independent of each other and independent of  $X(0)$ , then  $\mathcal{U}_1, \dots, \mathcal{U}_k$  are independent of each other in the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_0)$ . Therefore, from Lemma 4.8 below  $t_1 \sim \exp(\sum_k \alpha_k(x_0))$  in the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_0)$ . Let  $J_1 \in \{(k, l_k(1)) : 1 \leq k \leq n\}$  satisfies

$$t_1 = \frac{S_{J_1}}{\alpha_{J_1(1)}(x_0)},$$

where  $J_1(1)$  is the type of reaction. From Lemma 4.8  $t_1$  and  $J_1$  are independent in the probability space

$$(\Omega, \mathcal{F}, \mathbb{P}_0).$$

Moreover, for each  $1 \leq k \leq n$

$$\mathbb{P}_0(J_1 = (k, l_k(1))) = \frac{\alpha_k(x_0)}{\sum_j \alpha_j(x_0)}.$$

Hence,  $(J_1, t_1)$  has the same distribution as the reaction type and the first reaction time given in Dood-Gillespie SSA.

Suppose that the event  $\{J_1 = (j_1, l_{j_1}(1))\}$  has positive  $\mathbb{P}_0$ -probability. Given the event  $\{X_0 = x_0, J_1 = (j_1, l_{j_1}(1))\}$ , define  $x_1 := x_0 + v[j_1]$  and

$$\mathbb{P}_1(\cdot) := \mathbb{P}_0(\cdot | J_1 = (j_1, l_{j_1}(1))).$$

For now on we are restricted to the event  $\{X_0 = x_0, J_1 = (j_1, l_{j_1}(1))\}$ .

Now we are interested in finding the time  $t_2$  when the second reaction is going to happen. By (4.20) stoichiometric changes can not happen in the time interval  $[t_1, t_2)$ , then  $\forall 1 \leq k \leq n$

$$\int_0^{t_2} \alpha_k(X(s)) ds = \alpha_k(x_0)\tau_1 + \alpha_k(x_1)\tau_2,$$

where  $\tau_2 := t_2 - t_1$  and  $\tau_1 := t_1$ . Note that the second reaction is going to happen when the last integral above hits some event of the internal clock of some  $\mathcal{U}_k$ . More precisely,  $\tau_2$  must satisfy

$$\tau_2 = \min_{1 \leq k \leq n} \left\{ \frac{S_{k, l_k(2)} - \alpha_k(x_0)\tau_1}{\alpha_k(x_1)} \right\},$$

where  $l_k(2)$  is the first event of  $\mathcal{U}_k$  such that  $S_{k, l_k(2)} - \alpha_k(x_0)\tau_1 > 0$ . Observe that if  $J_1 = (j_1, l_{j_1}(1))$ , then a.s.  $\forall k \in \{1, \dots, n\}$  with  $k \neq j_1$

$$l_k(2) = l_k(1)$$

and

$$l_{j_1}(2) = l_{j_1}(1) + 1.$$



Hence,  $\tau_2$  satisfies

$$\tau_2 = \min \left\{ \frac{S_{k,1} - \alpha_k(x_0)\tau_1}{\alpha_k(x_1)}, \frac{S_{j_1,2} - \alpha_{j_1}(x_0)\tau_1}{\alpha_{j_1}(x_1)} : 1 \leq k \leq n, k \neq j_1 \right\}.$$

From Corollary 4.17, the set of random variables

$$\{S_{k,1} - \alpha_k(x_0)\tau_1 : 1 \leq k \leq n, k \neq j_1\}$$

are i.i.d. with distribution  $\exp(1)$  in the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_1)$ . Since

$$\mathbb{P}_1(\alpha_{j_1}(x_0)\tau_1 = S_{j_1,1}) = 1,$$

then the set of random variables

$$\{S_{k,1} - \alpha_k(x_0)\tau_1, S_{j_1,2} - \alpha_{j_1}(x_0)\tau_1 : 1 \leq k \leq n, k \neq j_1\}.$$

are i.i.d. with distribution  $\exp(1)$  in the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_1)$ . Therefore, from Lemma 4.8  $\tau_2 \sim \exp(\sum_k \alpha_k(x_1))$  in the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_1)$ . Moreover, if  $J_2 \in \{(k, l_k(2)) : 1 \leq k \leq n\}$  satisfies

$$\tau_2 = \frac{S_{J_2} - \alpha_{J_2(1)}(x_0)\tau_1}{\alpha_{J_2(1)}(x_1)},$$

then  $\tau_2$  and  $J_2$  are independent in  $(\Omega, \mathcal{F}, \mathbb{P}_1)$ ; also  $\forall 1 \leq k \leq n$

$$\mathbb{P}_1(J_2 = (k, l_k(2))) = \frac{\alpha_k(x_1)}{\sum_j \alpha_j(x_1)}.$$

Let  $a \in \mathbb{R}_0^+$  and note that Proposition (4.16) implies that  $\tau_2$  and  $J_2$  are independent and have the same distributions in the probability spaces

$$(\Omega, \mathcal{F}, \mathbb{P}_1), \quad (\Omega, \mathcal{F}, \mathbb{P}_0(\cdot | J_1 = (j_1, l_{j_1}(1)), \tau_1 > a)).$$

Since this is  $\forall a \in \mathbb{R}_0^+$ , we conclude that

$$(\tau_1, \tau_2, J_2) \text{ are independent in the probability space } (\Omega, \mathcal{F}, \mathbb{P}_1).$$

In summary, we have shown that  $(\tau_1, \tau_2, J_1, J_2)$  have the same stochastic attributes than the Doob-Gillespie's SSA.

We can summarize this procedure in the following algorithm.

**Algorithm 4.21.** (*Anderson-Kurtz's SSA*)

- (1) *Initialization:* Set  $x \leftarrow x_0$ ,  $Int \leftarrow 0 \in \mathbb{R}^n$ , and  $t \leftarrow 0$ .
- (2) *Generate  $n$  independent random numbers  $s_1, \dots, s_n$  with distribution  $\exp(1)$ , and define the vector  $S$  as  $S \leftarrow (s_1, \dots, s_n)$ .*
- (3) *Calculate  $\alpha_1(x), \dots, \alpha_n(x)$  and set*

$$\alpha \leftarrow (\alpha_1(x), \dots, \alpha_n(x)).$$

- (4) *Define  $B$  as*

$$B \leftarrow \{k \in \{1, \dots, n\} : \alpha_k \neq 0\}$$

- (5) If  $B \neq \emptyset$ , then set  $\tau = \min\{\frac{S_k - Int_k}{\alpha_k} : k \in B\}$ . Otherwise, the algorithm ends.
- (6) Fix  $K$  as the set  $K = \{k \in B : \tau = \frac{S_k - Int_k}{\alpha_k}\}$ .
- (7) Generate a set  $\{s_k : k \in K\}$  of independent random numbers with distribution  $\exp(1)$ . Update the vector  $S$  by modifying the entries

$$S_k \leftarrow S_k + s_k$$

for each  $k \in K$ .

- (8) Update  $Int$  as

$$Int \leftarrow Int + \tau\alpha.$$

- (9) Update  $x$  and  $t$  as

$$x \leftarrow x + \sum_{k \in K} v[k],$$

$$t \leftarrow t + \tau.$$

- (10) Return to step (3).

Moreover we get the following result.

**Theorem 4.22.** Consider a reaction system of  $n$  reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n$  and  $m$  chemical species  $S_1, \dots, S_m$ . For this chemical system, the following is true:

- (1) Anderson-Kurtz's SSA is equivalent to Doob-Gillespie's SSA .
- (2) Anderson-Kurtz's equation for this chemical system

$$X(t) = X(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(X(s)) ds \right) v[k]$$

has as unique solution the reaction jump process given in (4.1) and Section 2.2.

*Proof.* We proceed by induction. Note that the base of the induction step is given above when we generate  $\tau_1, \tau_2, J_1, J_2$ . Let  $q \in \mathbb{N}$   $q > 1$ . Now suppose that

$$\{J_1 = (j_1, c_1), \dots, J_{q-1} = (j_{q-1}, c_{q-1}), J_q = (j_q, c_q)\}$$

has positive  $\mathbb{P}_0$ -probability. For each  $1 \leq i \leq q$

$$\mathbb{P}_i(\cdot) = \mathbb{P}_{i-1}(\cdot | J_i = (j_i, c_i)),$$

$$x_i = x_{i-1} + v[j_i];$$

i.e., for each  $1 \leq i \leq q$

$$\mathbb{P}_i(\cdot) = \mathbb{P}_0(\cdot | J_1 = (j_1, c_1), \dots, J_i = (j_i, c_i)).$$

We take the following induction hypothesis: we assume that for each  $1 \leq i \leq q$ , the time of the  $i$ th reaction is given by  $t_i = \sum_{r=1}^i \tau_r$  where  $\tau_1, \dots, \tau_q$  satisfy:

(1) For each  $1 \leq i \leq q$ , the time step  $\tau_i$  is given by

$$\tau_i = \min_{1 \leq k \leq n} \left\{ \frac{S_{k, l_k(i)} - \left( \sum_{r=1}^{i-1} \alpha_k(x_{r-1}) \tau_r \right)}{\alpha_k(x_{i-1})} \right\}.$$

where  $l_k(i)$  is the first event of  $\mathcal{U}_k$  such that

$$S_{k, l_k(i)} - \left( \sum_{r=1}^{i-1} \alpha_k(x_{r-1}) \tau_r \right) > 0,$$

and  $\tau_i \sim \exp(\sum_{k=1}^n \alpha_k(x_{i-1}))$  in the probability space  $(\Omega; \mathcal{F}; \mathbb{P}_{i-1})$

(2)  $\tau_1, \dots, \tau_q, J_q$  are  $\mathbb{P}_{q-1}$ -independent.

(3) The following set is i.i.d. with distribution  $\exp(1)$  in the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_{q-1})$

$$\left\{ S_{k, l_k(q)} - \left( \sum_{r=1}^{q-1} \alpha_k(x_{r-1}) \tau_r \right) : 1 \leq k \leq n \right\}.$$

Let  $t_{q+1}$  be the time of the  $(q+1)$ th reaction and define  $\tau_{q+1} := t_{q+1} - t_q$ . From (4.20) for each  $1 \leq k \leq n$

$$\int_0^{t_{q+1}} \alpha_k(X(s)) ds = \sum_{r=1}^{q+1} \alpha_k(x_{r-1}) \tau_r,$$

which implies that  $\tau_{q+1}$  must satisfy

$$\tau_{q+1} = \min_{1 \leq k \leq n} \left\{ \frac{S_{k, l_k(q+1)} - \left( \sum_{r=1}^q \alpha_k(x_{r-1}) \tau_r \right)}{\alpha_k(x_q)} \right\},$$

where  $l_k(q+1)$  is the first event of  $\mathcal{U}_k$  such that

$$S_{k, l_k(q+1)} - \left( \sum_{r=1}^q \alpha_k(x_{r-1}) \tau_r \right) > 0.$$

From corollary (4.17) the set

$$\left\{ \left( S_{k, l_k(q)} - \sum_{r=1}^{q-1} \alpha_k(x_{r-1}) \tau_r \right) - \alpha_k(x_{q-1}) \tau_q : 1 \leq k \leq n \quad k \neq i_q \right\}$$

is i.i.d. with distribution  $\exp(1)$  in the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_q)$  and

$$\mathbb{P}_q \left( S_{j_q, l_{j_q}(q)} - \left( \sum_{r=1}^q \alpha_{j_q}(x_{r-1}) \tau_r \right) = 0 \right) = 1.$$

Since we are restricted to the event

$$\{J_1 = (j_1, c_1), \dots, J_{q-1} = (j_{q-1}, c_{q-1}), J_q = (j_q, c_q)\},$$

$\forall 1 \leq k \leq n \quad k \neq i_q,$

$$l_k(q+1) = l_k(q)$$

and

$$l_{j_q}(q+1) = l_{j_q}(q) + 1.$$

Then  $\tau_{q+1}$  satisfies  $\mathbb{P}_q - a.s.$

$$\tau_{q+1} = \min_{1 \leq k \leq n, k \neq j_q} \left\{ \frac{S_{k, l_k(q)} - (\sum_{r=1}^q \alpha_k(x_{r-1}) \tau_r)}{\alpha_k(x_q)}, \frac{S_{j_q, l_{j_q}(q)+1} - S_{j_q, l_{j_q}(q)}}{\alpha_{j_q}(x_q)} \right\},$$

$$\tau_{q+1} = \min_{1 \leq k \leq n, k \neq j_q} \left\{ \frac{\left( S_{k, l_k(q)} - \sum_{r=1}^{q-1} \alpha_k(x_{r-1}) \tau_r \right) - \alpha_k(x_{q-1}) \tau_q}{\alpha_k(x_q)}, \frac{S_{j_q, l_{j_q}(q)+1} - S_{j_q, l_{j_q}(q)}}{\alpha_{j_q}(x_q)} \right\}.$$

From lemma (4.8),  $\tau_{q+1} \sim \exp(\sum_k \alpha_k(x_q))$ , and if

$$J_{q+1} \in \{(k, l_k(q+1)) : 1 \leq k \leq n\}$$

is such that

$$\tau_{q+1} = \frac{S_{J_{q+1}} - \sum_{r=1}^q \alpha_{J_{q+1}(1)}(x_{r-1}) \tau_r}{\alpha_{J_{q+1}(1)}(x_q)},$$

then  $\tau_{q+1}$  and  $J_{q+1}$  are  $\mathbb{P}_q$ -independent and

$$\mathbb{P}_n(J_{q+1} = (k, l_k(q+1))) = \frac{\alpha_k(x_q)}{\sum_j \alpha_j(x_q)}.$$

Now we are going to show that  $\tau_1, \dots, \tau_q, \tau_{q+1}, J_{q+1}$  are  $\mathbb{P}_q$  independent. First, note from (2) above that  $\tau_1, \dots, \tau_q$  are  $\mathbb{P}_q$ -independent. Now let  $a_1, \dots, a_q$  be nonnegative real numbers. From Proposition (4.16)  $\tau_{q+1}, J_{q+1}$  are independent and have the same distributions with the following probabilities measures

$$\begin{aligned} & \mathbb{P}_0(\cdot | J_1=(j_1, c_1), J_2=(j_2, c_2), \dots, J_q=(j_q, c_q)) \\ & \mathbb{P}_0(\cdot | J_1=(j_1, c_1), \tau_1 > a_1, J_2=(j_2, c_2), \dots, J_q=(j_q, c_q)) \\ & \mathbb{P}_0(\cdot | J_1=(j_1, c_1), \tau_1 > a_1, J_2=(j_2, c_2), \tau_2 > a_2, J_3=(j_3, c_3), \dots, J_q=(j_q, c_q)) \\ & \vdots \\ & \mathbb{P}_0(\cdot | J_1=(j_1, c_1), \tau_1 > a_1, J_2=(j_2, c_2), \tau_2 > a_2, \dots, J_q=(j_q, c_q), \tau_q > a_q). \end{aligned}$$

Since this is for every  $a_1, \dots, a_q$  nonnegative real numbers and  $\tau_1, \dots, \tau_q$  are  $\mathbb{P}_q$ -independent, we conclude that

$$\tau_1, \dots, \tau_q, \tau_{q+1}, J_{q+1}$$

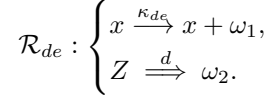
are  $\mathbb{P}_q$ -independent.  $\square$

Note that in point (7) of Algorithm 4.21, the probability that  $|K| > 1$  is equal to 0. Therefore, if we exclude point (2) of Algorithm 4.21, then every step of Anderson-Kurtz's SSA only generates one random variable. This is a big difference with Doob-Gillespie's SSA because every step of Doob-Gillespie's SSA generates two random variables (see Algorithm 4.4).

## 5. SIMULATION OF THE REACTION JUMP PROCESS WITH DELAYS

## 5.1. Anderson-Kurtz's Stochastic Simulation Algorithm with Delays (SSAD).

Now we study the simulation of the reaction jump process of a chemical system with delays. Consider again a chemical system of  $n$  chemical reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n$  and one chemical reaction with delay  $\mathcal{R}_{de}$



It is important to remember that for this system the stoichiometric changes not only depend on normal reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n, x \xrightarrow{\kappa_{de}} x + \omega_1$ , but they also rely on the decompositions of the  $Z$  complexes. Therefore, two questions arise: How can we simulate the times where the stoichiometry of the system will change? And, how can we know if the stoichiometric change is due to a normal reaction or to the decomposition of the excited complex  $Z$ ? We are going to show how Anderson-Kurtz's equation can answer these questions in a simple form.

As proposed in Section 3.1, the dynamics of this reaction system is given by the Anderson-Kurtz's equation

$$(5.1) \quad \begin{aligned} Y(t) = Y(0) &+ \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(Y(s)) ds \right) v[k] \\ &+ \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \omega_1 + \mathcal{U}_{de} \left( \int_0^{t-d} \alpha_{de}(Y(s)) ds \right) \omega_2. \end{aligned}$$

where every possible solution of (5.1) has to be of the form

$$(5.2) \quad Y(t) = \sum_{j=0}^{\infty} Y_j \chi_{[\hat{t}_j, \hat{t}_{j+1})}(t),$$

where the embedded chain  $\{Y_j\}_{j \in \mathbb{N}_0}$  satisfies  $Y_{j+1} = Y_j + w$ , for some

$$w \in \{v_1, \dots, v_n, \omega_1, \omega_2\};$$

and  $\{\hat{t}_j\}$  are the times where there is a stoichiometric change due to a chemical reaction or decomposition of a complex  $Z$ . For example one possible event that we could have is the following one:

$$\begin{array}{ccccccc} \mathcal{R}_{\star_1} & \mathcal{R}_{\star_2} & Z & \mathcal{R}_{\star_3} & Z & Z & \mathcal{R}_{\star_4}, \\ \hat{t}_1 & \hat{t}_2 & \hat{t}_3 & \hat{t}_4 & \hat{t}_5 & \hat{t}_6 & \hat{t}_7 \end{array}$$

where at time  $\hat{t}_1$  occurs a chemical reaction  $\mathcal{R}_{\star_1}$ , at time  $\hat{t}_2$  occurs a chemical reaction  $\mathcal{R}_{\star_2}$ , at time  $\hat{t}_3$  an excited complex  $Z$  has decomposed, etc.

Suppose that  $Y(0)$  and the Poisson processes  $\mathcal{U}_1, \dots, \mathcal{U}_n, \mathcal{U}_{de}$  are well adapted to the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ ; also suppose that  $Y(0)$  is independent of  $\mathcal{U}_1, \dots, \mathcal{U}_n, \mathcal{U}_{de}$ . Let  $y_0$  be some state of the system such that  $\mathbb{P}(Y(0)=y_0) > 0$ ,

$$\forall \star \in \{1, \dots, n, de\} \quad \{S_{\star, j}\}_j \quad \text{is the sequence of events of } \mathcal{U}_{\star}.$$

Note that  $\star$  is only a symbol, so that  $\mathcal{U}_\star$  means  $\mathcal{U}_{de}$  when  $\star=de$ . Remember that  $\mathcal{U}_1, \dots, \mathcal{U}_n, \mathcal{U}_{de}$  are independent unit-rate Poisson processes, therefore the sequences defined above are independent.

Following equations (5.1) and (5.2), if we are restricted to the event  $\{Y(0) = y_0\}$ , then  $\forall s \in [0, \hat{t}_1)$ ,  $Y(s) = y_0$  and  $\forall \star \in \{1, \dots, n, de\}$

$$\int_0^{\hat{t}_1} \alpha_\star(Y(s)) ds = \hat{t}_1 \alpha_\star(y_0);$$

and if at time  $\hat{t}_1$  occurs some reaction  $\mathcal{R}_{j_1}$ , then

$$\hat{t}_1 \alpha_{j_1}(y_0) = S_{j_1, l_{j_1}(1)} \iff \hat{t}_1 = \frac{S_{j_1, l_{j_1}(1)}}{\alpha_{j_1}(y_0)}$$

and  $\forall \star \in \{1, \dots, n, de\} \setminus \{j_1\}$

$$\hat{t}_1 \alpha_\star(y_0) < S_{\star, l_{j_1}(1)} \iff \hat{t}_1 < \frac{S_{\star, l_{j_1}(1)}}{\alpha_\star(y_0)}.$$

That is, the time  $\hat{t}_1$  satisfy

$$\hat{t}_1 = \min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_\star(1)}}{\alpha_\star(y_0)} \right\},$$

where  $l_\star(1)$  is the first event of  $\mathcal{U}_\star$  such that  $S_{\star, l_\star(1)} > 0$ . It follows that  $\hat{t}_1$  is a.s.

$$\hat{t}_1 = \min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, 1}}{\alpha_\star(y_0)} \right\},$$

which implies that  $\hat{t}_1 \sim \exp(\sum_\star \alpha_\star(y_0))$ . If

$$J_1 \in \{(\star, l_\star(1)) : \star \in \{1, \dots, n, de\}\}$$

is the index that occurs, then  $\hat{t}_1, J_1$  are  $\mathbb{P}(\cdot | Y(0) = y_0)$ -independent and

$$\mathbb{P}(J_1 = (j_1, l_{j_1}(1)) | Y(0) = y_0) = \frac{\alpha_{j_1}(y_0)}{\sum_\star \alpha_\star(y_0)},$$

where  $j_1 \in \{1, \dots, n, de\}$ . Note that the stoichiometric change at  $\hat{t}_1$  is due to a chemical reaction, which is what is expected when there are no molecules of excited complex  $Z$  in the system (which is one of the initial conditions).

Given the event  $\{Y(0) = y_0, J_1 = (j_1, l_{j_1}(1))\}$  of positive  $\mathbb{P}$ -probability, we are interested in the time  $\hat{t}_2$  where the second stoichiometric change will occur. Here we need to deal with two cases:  $j_1 \neq de$  and  $j_1 = de$ .

For case  $j_1 \neq de$ , the system has not formed excited complexes at time  $\hat{t}_1$ , in terms of equation (5.1)

$$\int_0^{\hat{t}_1} \alpha_{de}(Y(s)) ds < S_{de, 1},$$

which implies that

$$\mathcal{U}_{de} \left( \int_0^{\hat{t}_1+d-d} \alpha_{de}(Y(s))ds \right) = \mathcal{U}_{de} \left( \int_0^{\hat{t}_1} \alpha_{de}(Y(s))ds \right) = 0;$$

that is, the next stoichiometric change is due to a chemical reaction. Let  $y_1 = y_0 + v[j_1]$  and note that (5.2) implies that for  $\star \in \{1, \dots, n, de\}$

$$\int_0^{\hat{t}_2} \alpha_\star(Y(s))ds = \alpha_\star(y_0)\hat{\tau}_1 + \alpha_\star(y_1)\hat{\tau}_2,$$

where  $\hat{\tau}_2 := \hat{t}_2 - \hat{t}_1$  and  $\hat{\tau}_1 := \hat{t}_1$ . Then

$$\hat{\tau}_2 = \min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_\star(2)} - \alpha_\star(y_0)\tau_1}{\alpha_\star(y_1)} \right\},$$

where  $l_\star(2)$  is the first event of  $\mathcal{U}_\star$  such that  $S_{\star, l_\star(2)} - \alpha_\star(y_0)\tau_1 > 0$ . From Section 4.2 we know that  $\hat{\tau}_2 \sim \exp(\sum_\star \alpha_\star(y_1))$  with the probability measure

$$\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (j_1, l_{j_1}(1))).$$

Now then, if  $J_2$  is the index that occurs, then  $\tau_2$  and  $J_2$  are

$$\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (j_1, l_{j_1}(1)))\text{-independent}$$

and  $J_2$  has the same distribution as in Section 4.2.

Now suppose we got case  $J_1 = (de, l_{de}(1))$ . Thus given the event

$$\{Y(0) = y_0, J_1 = (de, l_{de}(1))\},$$

with probability 1, an stoichiometric change will occur at time  $\hat{t}_1 + d$  because

$$\mathcal{U}_{de} \left( \int_0^{\hat{t}_1+d-d} \alpha_{de}(Y(s))ds \right) = \mathcal{U}_{de} \left( \int_0^{\hat{t}_1} \alpha_{de}(Y(s))ds \right) = 1;$$

that is,  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)))$  -a.s. exists  $\hat{t}_i$  in the sequence  $\{\hat{t}_j\}$  such that  $\hat{t}_i = \hat{t}_1 + d$ .

At this point, the second stoichiometric change may be due to the decomposition of the first excited complex  $Z$  or to a chemical reaction, that is  $\hat{t}_2 = \hat{t}_1 + d$  or

$$\hat{t}_2 = \hat{t}_1 + \min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_\star(2)} - \alpha_\star(y_0)\hat{\tau}_1}{\alpha_\star(y_1)} \right\}.$$

Therefore, if  $\hat{t}_2 = \hat{t}_1 + \hat{\tau}_2$ , then

$$\hat{\tau}_2 = \min \left\{ \min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_\star(2)} - \alpha_\star(y_0)\hat{t}u_1}{\alpha_\star(y_1)} \right\}, d \right\},$$

where  $\hat{\tau}_1 = \hat{t}_1$ ,  $y_1 = y_0 + \omega_1$  and from Section 4.2

$$\min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_\star(2)} - \alpha_\star(y_0)\hat{\tau}_1}{\alpha_\star(y_1)} \right\} \text{ has distribution } \exp \left( \sum_\star \alpha_\star(y_1) \right)$$

and is independent of  $\hat{\tau}_1$  with the probability measure  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)))$ .  
Let

$$\{J_2 = Z\} = \{\hat{\tau}_2 = d\};$$

that is, the first complex formed at time  $\hat{t}_1$  has decomposed at time  $\hat{t}_2$ . Note that we have the following event equality

$$(5.3) \quad \{J_2 = Z\} = \{S_{\star, l_{\star}(2)} - \alpha_{\star}(y_0)\hat{\tau}_1 \geq \alpha_{\star}(y_1)\hat{\tau}_2 : \star \in \{1, \dots, n, de\}\}.$$

Since for each  $\star \in \{1, \dots, n, de\}$

$$S_{\star, l_{\star}(2)} - \alpha_{\star}(y_0)\hat{\tau}_1 \quad \text{has distribution} \quad \exp(1)$$

and are independent with the probability measure  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)))$ . Then, from the memoryless property of the exponential distribution, the following set

$$\{S_{\star, l_{\star}(2)} - \alpha_{\star}(y_0)\hat{\tau}_1 - \alpha_{\star}(y_1)\hat{\tau}_2 : \star \in \{1, \dots, n, de\}\}$$

is i.i.d. with  $\exp(1)$  distribution under the probability measure

$$\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z).$$

Observe that given event  $\{Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z\}$ , the equations (5.2) and (5.1) implies that for each  $\star \in \{1, \dots, n, de\}$ ,

$$\int_0^{\hat{t}_3} \alpha_{\star}(Y(s))ds = \alpha_{\star}(y_0)\hat{\tau}_1 + \alpha_{\star}(y_1)\hat{\tau}_2 + \alpha_{\star}(y_2)\hat{\tau}_3,$$

where  $\hat{\tau}_3 = \hat{t}_3 - \hat{t}_2$ ,  $y_1 = y_0 + \omega_1$  and  $y_2 = y_1 + \omega_2$ . Since at time  $\hat{t}_2$  there are not molecules of excited complex  $Z$ , then the next stoichiometric change will be given by a chemical reaction with  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z)$ -probability 1. That is,  $\hat{\tau}_3$  satisfies

$$\hat{\tau}_3 = \min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_{\star}(3)} - \alpha_{\star}(y_0)\hat{\tau}_1 - \alpha_{\star}(y_1)\hat{\tau}_2}{\alpha_{\star}(y_2)} \right\}$$

where  $l_{\star}(3)$  is the first time event of  $\mathcal{U}_{\star}$  such that

$$S_{\star, l_{\star}(3)} - \alpha_{\star}(y_0)\tau_1 - \alpha_{\star}(y_1)\tau_2 > 0.$$

From (5.3)  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z)$ -a.s. we get

$$\forall \star \in \{1, \dots, n, de\} \quad l_{\star}(3) = l_{\star}(2).$$

This implies that  $\hat{\tau}_3$  satisfies  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z)$ -a.s.

$$\hat{\tau}_3 = \min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_{\star}(2)} - \alpha_{\star}(y_0)\hat{\tau}_1 - \alpha_{\star}(y_1)\hat{\tau}_2}{\alpha_{\star}(y_2)} \right\},$$



which implies that  $\hat{\tau}_3$  has  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z)$ -distribution

$$\hat{\tau}_3 \sim \exp\left(\sum_{\star} \alpha_{\star}(y_2)\right).$$

Although,  $\hat{\tau}_3$  follows an exponential distribution as above, it is important to note that  $\hat{\tau}_3$  and

$$\min_{\star \in \{1, \dots, n, de\}} \left\{ \frac{S_{\star, l_{\star}(2)} - \alpha_{\star}(y_0)\tau_1}{\alpha_{\star}(y_1)} \right\}$$

are not  $\mathbb{P}(\cdot | Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z)$ - independent. In other words, given the event

$$\{Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = Z\},$$

the competition between the chemical reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n, x \xrightarrow{\kappa_{de}} x + \omega_1$  at time  $\hat{t}_1$  is no longer independent of the competition on these chemical reactions at time  $\hat{t}_2$ .

The procedure described above can be summarized in the following algorithm. The information of the reaction with delay  $\mathcal{R}_{de}$  is stored in the index  $n+1$ .

**Algorithm 5.4.** (*Anderson-Kurtz's SSAD*)

- (1) *Initialization:* Set  $y \leftarrow y_0$ ,  $Int \leftarrow 0 \in \mathbb{R}^{n+1}$ ,  $t \leftarrow 0$  and  $D = \{\infty\}$ .
- (2) *Generate*  $n+1$  independent and  $\exp(1)$  random numbers  $s_1, \dots, s_n, s_{n+1}$  and set  $S$  as  $S \leftarrow (s_1, \dots, s_n, s_{n+1})$ .
- (3) *Calculate*  $\alpha_1(y), \dots, \alpha_n(y), \alpha_{n+1}(y)$  and set

$$\alpha \leftarrow (\alpha_1(y), \dots, \alpha_n(y), \alpha_{n+1}(y)).$$

- (4) *Set*  $B \subset \{1, \dots, n+1\}$  as

$$B \leftarrow \{1 \leq k \leq n+1 : \alpha_k \neq 0\}.$$

- (5) *If*  $B \neq \emptyset$ , then set  $\hat{\tau} = \min \left( \left\{ \frac{S_k - Int_k}{\alpha_k} : k \in B \right\} \cup D \right)$ . *Otherwise, the algorithm ends.*
- (6) *If*  $\hat{\tau} \in D$ , *update*  $y$  by  $y \leftarrow y + \omega_2$  and  $D$  by

$$D = D \setminus \{\hat{\tau}\}.$$

*Otherwise go to the next step.*

- (7) *Update*  $D$  as

$$D \leftarrow D - \hat{\tau}.$$

- (8) *Set*  $K \subset \{1, \dots, n+1\}$  as  $K = \left\{ k \in B : \tau = \frac{S_k - Int_k}{\alpha_k} \right\}$ . *If*  $K = \emptyset$ , then go to step (11). *Otherwise update*  $y$  as

$$y \leftarrow y + \sum_{k \in K} v[k], \text{ and go to the next step.}$$

- (9) Generate a set  $\{s_k : k \in K\}$  of independent random numbers with distribution  $\exp(1)$ . Update  $S$  entry by entry

$$S_k \leftarrow S_k + s_k$$

for each  $k \in K$ .

- (10) If  $n+1 \in K$ , then update  $D$  as  $D \leftarrow D \cup \{d\}$ . Otherwise go to the next step.  
 (11) Update  $Int$  as

$$Int \leftarrow Int + \hat{\tau}\alpha.$$

- (12) Update the time  $t$  as

$$t \leftarrow t + \hat{\tau}.$$

- (13) Return to step (3).

Now we make some observations to the previous Algorithm 5.4:

- (1) The information of the reaction with delay  $\mathcal{R}_{de}$  is stored in the index  $n+1$ .
- (2)  $D - \hat{\tau} = \{a - \hat{\tau} : a \in D\}$ .
- (3) With probability zero  $\hat{\tau} \in D$  and  $K \neq \emptyset$ .
- (4) With probability zero  $|K| > 1$ .

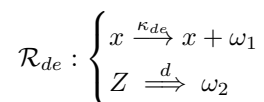
It is important to note that in general, given the information of the first  $q$  reactions ( $q \in \mathbb{N}$ ), the set of delays at time  $\hat{t}_q$ , say  $D_q$ , are random variables. For example, suppose that we have the following event

$$\{Y(0) = y_0, J_1 = (de, l_{de}(1)), J_2 = (de, l_{de}(2))\},$$

that is, at time  $\hat{t}_1$  occurs a reaction with delay and at time  $\hat{t}_2$  occurs another reaction with delay. Therefore, at time  $\hat{t}_2$  the set of delays satisfies

$$D_2 = \{d - \hat{\tau}_2, d\}.$$

**5.2. Barrio's SSAD.** As we said in the introduction the Anderson-Kurtz's algorithm was not the first algorithm designed to simulate chemical systems with delays. Actually, there are two algorithms previously designed for that purpose: Cai's algorithm and Barrio's algorithm. In this section we introduce the standard algorithm to simulate the reaction jump process of chemical system with delays; this algorithm is due to Barrio; see for example [3] and [5]. For the sake of simplicity here we only going to deal with the chemical system of  $n$  chemical reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n$  and one reaction with delay  $\mathcal{R}_{de}$



Like Anderson-Kurtz's SSAD, Barrio's SSAD deals with the problem of the simulation of the times where the stoichiometry of the system will change and if this stoichiometric change is due to a chemical reaction or to a decomposition of a complex  $Z$ .

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space; suppose that we know how the reaction jump process is given in the time interval  $[0, t)$ ; that is, we know  $\{Y(s)\}_{s \in [0, t)}$  and suppose that  $\sigma(Y(s) : s \in [0, t)) \subset \mathcal{F}$ . Suppose that at time  $t$  the reaction jump process satisfies  $Y(t)=y$  and at times

$$t+d_1, \dots, t+d_l$$

an excited complex  $Z$  will decompose. Then, if at time  $t+\hat{\tau}$  occurs the next stoichiometric change, Barrio's proposal is to calculate the waiting time  $\hat{\tau}$  as follows:

$$(5.5) \quad \hat{\tau} := \min\{\tau, d_1, \dots, d_l\},$$

where  $\tau$  has distribution  $\exp\left(\sum_{\star \in \{1, \dots, n, de\}} \alpha_{\star}(y)\right)$  and is independent of  $\sigma(Y(s) : s \in [0, t))$ , everything under the probability measure  $\mathbb{P}(\cdot | Y(t) = y)$ . That is,  $\tau$  satisfies

$$\mathbb{P}(\tau > a | \sigma(Y(s) : s \in [0, t)), Y(t) = y) = \mathbb{P}(\tau > a | Y(t) = y) = \exp(-\alpha_{sum}(y)a),$$

where  $\alpha_{sum}(y) = \sum_{\star \in \{1, \dots, n, de\}} \alpha_{\star}(y)$ .

If  $\hat{\tau} = \tau$ , then Barrio's SSAD updates the state of the system as

$$y \leftarrow y + v[I],$$

where  $I$  satisfies:  $\forall \star \in \{1, \dots, n, de\}$

$$(5.6) \quad \mathbb{P}(I = \star | \sigma(Y(s) : s \in [0, t)), Y(t) = y) = \frac{\alpha_{\star}(y)}{\alpha_{sum}(y)}$$

and  $I$  is  $\mathbb{P}(\cdot | Y(t) = y)$ -independent of  $\tau$ . Here  $v[de] := \omega_1$ .

If  $\hat{\tau} \in \{d_1, \dots, d_l\}$ , then Barrio's SSAD updates the system as

$$y \leftarrow y + \omega_2.$$

Therefore, Barrio's proposal for the reaction jump process is given by:

$$\forall s \in [t, t+\hat{\tau}) \quad Y(s) = y$$

and according to (5.5),

$$Y(t+\hat{\tau}) = \begin{cases} y+w & w \in \{v[k], \omega_1 : 1 \leq k \leq n\} \quad \text{if } \hat{\tau} = \tau, \\ y+\omega_2 & \text{if } \hat{\tau} \in \{d_1, \dots, d_l\}. \end{cases}$$

This procedure is summarized in the following algorithm. All the information of the reaction  $\mathcal{R}_{de}$  is in the index  $n+1$ .

**Algorithm 5.7.** (*Barrio's SSAD*)

- (1) *Initialization:* Set  $y \leftarrow y_0$ ,  $t \leftarrow 0$ , and  $D = \{\infty\}$ .
- (2) *Calculate*  $\alpha_1(y), \dots, \alpha_n(y), \alpha_{n+1}(y)$  *and define*

$$\alpha \leftarrow (\alpha_1(y), \dots, \alpha_{n+1}(y)).$$

(3) Generate  $\tau \sim \exp(\sum_{k=1}^{n+1} \alpha_k)$ , and set

$$\hat{\tau} = \min(\{\tau\} \cup D).$$

(4) Update the time  $t \leftarrow t + \hat{\tau}$ .

(5) If  $\hat{\tau} \in D$ , then update the state of the system as

$$y \leftarrow y + \omega_2$$

and update  $D$  as

$$D \leftarrow (D \setminus \{\hat{\tau}\}) - \{\hat{\tau}\} \quad \text{and return to step (2)}.$$

Otherwise go to the next step.

(6) Update  $D$  as

$$D \leftarrow D - \{\hat{\tau}\}$$

(7) Generate  $u_1 \sim U(0, 1)$ . Set  $i$  to be the smallest integer satisfying

$$\sum_{k=1}^{i-1} \alpha_k \leq u_1 \sum_k \alpha_k < \sum_{k=1}^i \alpha_k.$$

Update the state of the system as

$$y \leftarrow y + v[i].$$

(8) If  $i=n+1$ , update  $D$  as

$$D \leftarrow D \cup \{d\}.$$

(9) Return to step (2).

It is clear that Barrio's SSAD generates the sequence of times between stoichiometric changes  $\{\hat{\tau}_k\}_{k \in \mathbb{N}}$  and the sequence  $\{J_k\}_{k \in \mathbb{N}}$  where each

$$J_k \in \{I_k, Z\}$$

gives us the information on whether the stoichiometric change is due to a reaction ( $J_k=I_k$ ) or the decomposition of a complex ( $J_k=Z$ ) at time

$$\hat{t}_k := \sum_{j=1}^k \hat{\tau}_j.$$

Therefore, if at time  $\hat{t}_q$  we know the variables  $Y(0), J_1, \dots, J_q$  and the set of delays  $D$ , then we know from (5.6) how  $I_{q+1}$  is distributed and from equation (5.5)

$$\hat{\tau}_{q+1} = \min\{\tau_{q+1} \cup D\}.$$

where

$$\tau_{q+1} \sim \exp\left(Y(0) + \sum_{i=1}^q v[J_i]\right)$$

and  $v[Z] := \omega_2$ . In this sense, Barrio's SSAD is a clear modification of Doob-Gillepie's SSA (see for example algorithm (4.13)).

Now then, following the definition of  $\hat{\tau}$  in (5.5), we can interpret Barrio's SSAD in terms of a competition between chemical reactions and the decomposition of a complex  $Z$ . From Section 4.1, we know that  $\tau$  in (5.5) has an exponential distribution with parameter equal to the sum of the propensities and is independent of the past of the reaction jump process, which means that different chemical reactions are competing to occur first. With the definition of  $\hat{\tau}$  as the minimum of  $\tau$  and the delays in  $D$ , this competition now involves not only the reactions but also the delays.

Here we emphasize the fact that in every step of Barrio's SSAD  $\tau$  is independent of the past of the reaction jump process. For example, if we know  $Y(0), J_1, \dots, J_q$ , then until time  $\hat{t}_q$  Barrio's SSAD has generated  $\tau_1, \dots, \tau_q$  random variables which satisfy

$$\tau_k \sim \exp\left(Y(0) + \sum_{i=1}^k v[J_i]\right)$$

and are  $\mathbb{P}(\cdot | Y(0), J_1, \dots, J_q)$ -independent. This means that Barrio's SSAD does an important assumption in terms of the competition of the reactions; this assumption is that the competitions of the chemical reactions at times  $\hat{t}_1, \dots, \hat{t}_q$  are independent.

It is clear from Algorithm 5.7 that the reaction jump process that Barrio's SSAD introduce has the following form

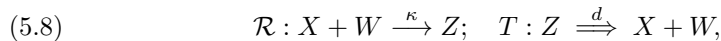
$$\forall t > 0 \quad Y(t) = \sum_{j=0}^{\infty} Y_j \chi_{[\hat{t}_j, \hat{t}_{j+1})}(t)$$

where  $\hat{t}_0 := 0$ ,  $Y_0 := Y(0)$ , and  $Y_{j+1} := Y_j + w$  for  $w \in \{v[k], \omega_1, \omega_2 : 1 \leq k \leq n\}$ . Note that this process has a similar structure as the solution of the Anderson-Kurtz's equation for this chemical system with delays (see Section 5.1). Thus, a natural question arises: Is the reaction jump process produced by Barrio's SSAD a solution of the Anderson and Kurtz equation? In other words, Do the sequences  $\{Y_j\}$  and  $\{\hat{t}_j\}$  that Barrio's SSAD generates have the same distribution as those that are solutions of the Anderson-Kurtz's equation? Unfortunately we still don't have a general answer to these questions, and we have to continue with this investigation in the Doctoral studies.

We could try to answer these questions by comparing Anderson-Kurtz's SSAD and Barrio's SSAD, but as the following example will show, Barrio's SSAD is different from Anderson-Kurtz's SSAD, and Barrio's SSAD generates solutions of the Anderson-Kurtz's equation up to time  $\hat{t}_4$ .

We will do these comparisons using a simple system.

**5.3. Example.** Consider the reaction system of one reaction  $\mathcal{R}$  and a physical transition  $T$  with delay  $d$



and initial state  $Y(0)=(2, 2)^T$ . It is clear that the set of possible states is given by

$$\mathbb{X} = \{(2, 2)^T, (1, 1)^T, (0, 0)^T\}$$

Therefore, if we know the number of molecules of  $X$  at time  $t$  (let's use notation  $x(t)$ ), then we know the state of the system at time  $t$ . Thus, we are going to restrict our attention to the dynamics of the number of molecules of  $X$ . Before continuing with the simulation of this chemical system, we introduce the notation  $(\hat{t}_k, x_k; d_1, \dots, d_l)$  to indicate that at time  $\hat{t}_k$  there are  $x_k$  molecules of  $X$  and  $l$  excited complexes  $Z$  that will decompose at times  $\hat{t}_k + d_1, \dots, \hat{t}_k + d_l$ .

We firstly present Barrio's SSAD for the chemical system (5.8). Following Algorithm 5.7 at the initial time  $\hat{t}_0$  we set  $D=\{\infty\}$  and

$$\hat{\tau}_1 = \min(D \cup \{\tau_1\}),$$

where  $\tau_1 \sim \exp(4\kappa)$ . It is clear that  $\hat{\tau}_1 = \tau_1$ , and the system does a transition, with probability 1, from  $(\hat{t}_0, 2)$  to  $(\hat{t}_1, 1; d)$ , that is Barrio's SSAD set  $\hat{t}_1 = \hat{t}_0 + \hat{\tau}_1$

$$D = \{d, \infty\} \quad x(\hat{t}_1) = 1.$$

Note that the event  $\{J_1 = \mathcal{R}\}$  has probability 1, thus

$$\mathbb{P} = \mathbb{P}(\cdot | J_1 = \mathcal{R}).$$

At time  $\hat{t}_1$  Barrio's SSAD define

$$\hat{\tau}_2 = \min(D \cup \{\tau_2\}) = \min(d, \tau_2)$$

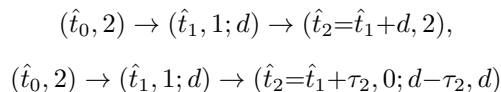
where  $\tau_2 \sim \exp(\kappa)$  and is  $\mathbb{P}$ -independent of  $\tau_1$ . If  $\hat{\tau}_2 = d$ , then at time  $\hat{t}_2 = \hat{t}_1 + d$  the next stoichiometric change is due to the decomposition of the only excited complex  $Z$  formed at time  $\hat{t}_1$ . This implies that Barrio's SSAD update  $D = \{\infty\}$  and  $x(\hat{t}_2) = 2$ . In the other case, at time  $\hat{t}_2 = \hat{t}_1 + \tau_2$  the next stoichiometric change is due to a new chemical reaction and Barrio's SSAD update  $D = \{d - \tau_2, d, \infty\}$  and  $x(\hat{t}_2) = 0$ . Note that

$$\mathbb{P}(J_1 = \mathcal{R}, J_2 = Z) = \mathbb{P}(\hat{\tau}_2 = d | J_1 = \mathcal{R}) \mathbb{P}(J_1 = \mathcal{R}) = \mathbb{P}(\tau_2 > d) = \exp(-\kappa d)$$

and

$$\mathbb{P}(J_1 = \mathcal{R}, J_2 = \mathcal{R}) = \mathbb{P}(\hat{\tau}_2 = \tau_2 | J_1 = \mathcal{R}) \mathbb{P}(J_1 = \mathcal{R}) = \mathbb{P}(\tau_2 \leq d) = 1 - \exp(-\kappa d).$$

This means that the following transitions



have probabilities  $\exp(-\kappa d)$  and  $1 - \exp(-\kappa d)$ , respectively.

Suppose that  $\hat{\tau}_2 = d$ ; then as we say before  $D = \{\infty\}$ ,  $x(\hat{t}_2) = 2$ . Thus, Barrio's SSAD continues the simulation by setting

$$\hat{\tau}_3 = \min(D \cup \{\tau_3\}) = \tau_3,$$

where  $\tau_3 \sim \exp(4\kappa)$  and  $\mathbb{P}(|J_1 = \mathcal{R}, J_2 = Z)$  – independent of  $\tau_1$  and  $\tau_2$ . This implies that with  $\mathbb{P}(|J_1 = \mathcal{R}, J_2 = Z)$  – probability 1 the system does the following transition

$$(\hat{t}_2 = \hat{t}_1 + d, 2) \rightarrow (\hat{t}_3 = \hat{t}_2 + \tau_3, 1, d)$$

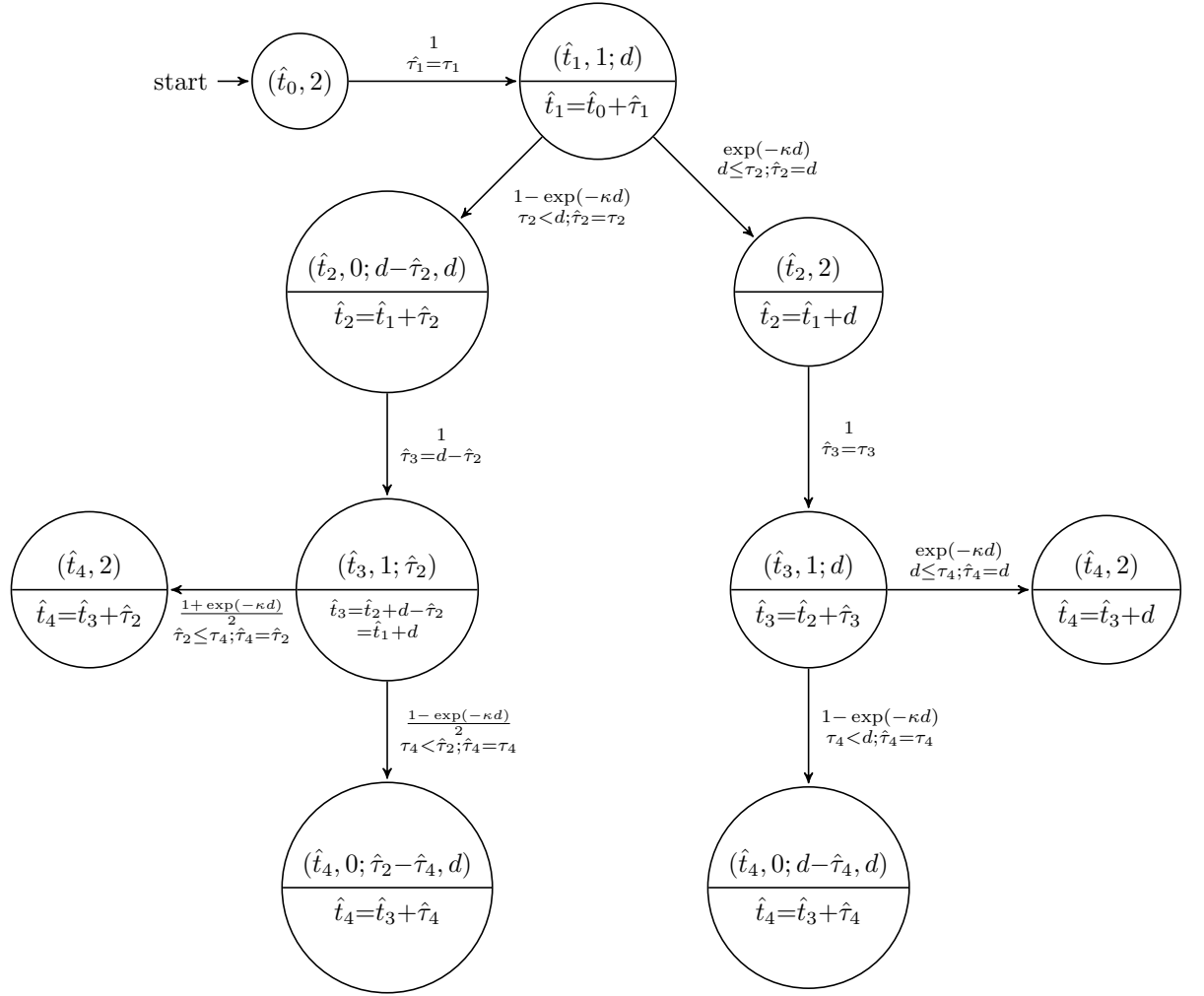
and the probability of the the following transitions

$$(\hat{t}_0, 2) \rightarrow (\hat{t}_1, 1; d) \rightarrow (\hat{t}_2 = \hat{t}_1 + d, 2) \rightarrow (\hat{t}_3 = \hat{t}_2 + \tau_3, 1, d)$$

is given by

$$\mathbb{P}(J_1 = \mathcal{R}, J_2 = Z, J_3 = \mathcal{R}) = \mathbb{P}(J_3 = \mathcal{R} | J_1 = \mathcal{R}, J_2 = Z) \mathbb{P}(J_2 = Z | J_1 = \mathcal{R}) \mathbb{P}(J_1 = \mathcal{R}) = 1 \times \exp(-\kappa d) \times 1.$$

The following diagram shows some of the paths that Barrio's SSAD can take for this system.



The label  $\left(\frac{p}{\hat{\tau}}\right)$  on the arrows indicates the probability  $p$  of taking the particular transition and how the waiting time  $\hat{\tau}$  is calculated. For instance, if we want to know

$$\mathbb{P}(J_1=\mathcal{R}, J_2=\mathcal{R}, J_3=\mathcal{R}, J_4=\mathcal{Z}),$$

then following the diagram we know that

$$\mathbb{P}(J_1=\mathcal{R})=1, \quad \mathbb{P}(J_2=\mathcal{R}|J_1=\mathcal{R})=1-\exp(-\kappa d), \quad \mathbb{P}(J_3=\mathcal{R}|J_1=\mathcal{R}, J_2=\mathcal{R})=1$$

and

$$\mathbb{P}(J_4=\mathcal{Z}|J_1=\mathcal{R}, J_2=\mathcal{R}, J_3=\mathcal{R})=\frac{1+\exp(-\kappa d)}{2}.$$

Therefore

$$\mathbb{P}(J_1=\mathcal{R}, J_2=\mathcal{R}, J_3=\mathcal{R}, J_4=\mathcal{Z})=(1-\exp(-\kappa d)) \times \frac{1+\exp(-\kappa d)}{2}.$$



Now we make the simulation that Anderson-Kurtz's SSAD generates for chemical system (5.8). Recall from Chapter 3 that the corresponding Anderson-Kurtz's equation for the system (5.8) is given by

$$\begin{pmatrix} x(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^t \kappa(xw)(s) ds \right) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^{t-d} \kappa(xw)(s) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Here we let the series  $\{S_k\}_{k \in \mathbb{N}}$  be the sequence of events of the unit-rate Poisson process  $\mathcal{U}_1$ .

Following Anderson-Kurtz's SSAD, at time  $\hat{t}_0=0$ ,  $D=\{\infty\}$ ,  $Int=0$ ,  $S=S_1$  and  $\hat{\tau}_1$  is given by

$$\hat{\tau}_1 = \min \left( \left\{ \frac{S-Int}{4\kappa} \right\} \cup D \right) = \frac{S_1}{4\kappa}.$$

Since  $S_1 \sim \exp(1)$ ,  $\hat{\tau}_1 \sim \exp(4\kappa)$ ; and because we have only one reaction the event  $\{J_1=\mathcal{R}\}$  has  $\mathbb{P}$ -probability 1.

At time  $\hat{t}_1=\hat{t}_0+\hat{\tau}_1$ , Anderson-Kurtz's SSAD updates  $D$ ,  $Int$ ,  $S$  and  $x(\hat{t}_1)$  as

$$D=\{d, \infty\}, \quad Int=4\kappa\hat{\tau}_1, \quad S=S_2 \quad x(\hat{t}_1)=1$$

and set  $\hat{\tau}_2$  as

$$\hat{\tau}_2 = \min \left( \left\{ \frac{S-Int}{\kappa} \right\} \cup D \right) = \min \left( \frac{S_2-4\kappa\hat{\tau}_1}{\kappa}, d \right) = \min \left( \frac{S_2-S_1}{\kappa}, d \right).$$

Note that  $\frac{S_2-S_1}{\kappa} \sim \exp(\kappa)$  and is  $\mathbb{P}(|J_1=\mathcal{R})$ -independent of  $\hat{\tau}_1$ . Therefore,  $\hat{\tau}_2$  has the same probabilistic properties as in Barrio's SSAD. Now we need to deal with two cases: if  $\hat{\tau}_2=\frac{S_2-S_1}{\kappa}$ , then Anderson-Kurtz's SSAD updates  $D$ ,  $Int$ ,  $S$  and  $x(\hat{t}_2)$  as

$$D=\{d-\hat{\tau}_2, d, \infty\}, \quad Int=4\kappa\hat{\tau}_1+\kappa\hat{\tau}_2=S_2, \quad S=S_3, \quad x(\hat{t}_2)=0.$$

In this case, at time  $\hat{t}_2$  the propensity of reaction  $\mathcal{R}$  is equal to zero; this implies that with  $\mathbb{P}(|J_1=\mathcal{R}, J_2=\mathcal{R})$ -probability 1 Anderson-Kurtz's SSAD set  $\hat{\tau}_3=d-\hat{\tau}_2$ . Observe that for this case  $\hat{\tau}_3$  has the same probabilistic properties as in Barrio's SSAD. At time  $\hat{t}_3$  Anderson-Kurtz's algorithm only updates  $D$ ,  $Int$  and  $x(\hat{t}_3)$  to

$$D=\{\hat{\tau}_2, \infty\}, \quad Int=4\kappa\hat{\tau}_1+\kappa\hat{\tau}_2+0\hat{\tau}_3=S_2, \quad x(\hat{t}_3)=1,$$

but does not update  $S$ ; that is for this cases  $S=S_3$ . This means that  $\hat{\tau}_4$  is set as

$$\hat{\tau}_4 = \min \left( \left\{ \frac{S-Int}{\kappa} \cup D \right\} \right) = \min \left( \frac{S_3-S_2}{\kappa}, \hat{\tau}_2 \right),$$

where  $\frac{S_3-S_2}{\kappa} \sim \exp(\kappa)$  and is independent of  $\hat{\tau}_2$ . Note that given the event

$$\{J_1=\mathcal{R}, J_2=\mathcal{R}, J_3=Z\}$$

the waiting time  $\hat{\tau}_4$  has the same probabilistic properties as in Barrio's SSAD.

Now suppose that  $\hat{\tau}_2=d$ ; then at time  $\hat{t}_2$  Anderson-Kurtz's SSAD only updates  $D$ ,

$Int$  and  $x(\hat{t}_2)$  to

$$D=\{\infty\}, \quad Int=4\kappa\hat{\tau}_1+\kappa d, \quad x(\hat{t}_2)=2,$$

but does not update  $S$ ; that is for this cases  $S=S_2$ . Then Anderson-Kurtz's SSAD set  $\hat{\tau}_3$  as

$$\hat{\tau}_3=\min\left(\left\{\frac{S-Int}{4\kappa}\right\}\cup D\right)=\frac{S_2-4\kappa\hat{\tau}_1-\kappa d}{4\kappa}=\frac{S_2-S_1-\kappa d}{4\kappa}.$$

Recall that in the analogue case Barrio's SSAD set  $\hat{\tau}_2=\min(\tau_2, d)$  and  $\hat{\tau}_3=\tau_3$ , where  $\tau_3$  and  $\tau_2$  are  $\mathbb{P}(|J_1=\mathcal{R}, J_2=Z)$ -independent. We can see that this is not the case for Anderson and Kurtz SSAD because  $\frac{S_2-S_1}{\kappa}$  and  $\frac{S_2-S_1-\kappa d}{4\kappa}$  are not  $\mathbb{P}(|J_1=\mathcal{R}, J_2=Z)$ -independent.

Despite this difference, given the event  $\{J_1=\mathcal{R}, J_2=Z\}$  the term  $\frac{S_2-S_1-\kappa d}{4\kappa}$  satisfies  $\forall a > 0$ ,

$$\begin{aligned} &\mathbb{P}\left(\frac{S_2-S_1-\kappa d}{4\kappa} > a \mid J_1=\mathcal{R}, J_2=Z\right) \\ &= \mathbb{P}\left(\frac{S_2-S_1-\kappa d}{4\kappa} > a \mid \hat{\tau}_2=d\right) \\ &= \mathbb{P}\left(\frac{S_2-S_1-\kappa d}{4\kappa} > a \mid \frac{S_2-S_1}{\kappa} > d\right) \end{aligned}$$

where

$$\mathbb{P}\left(\frac{S_2-S_1-\kappa d}{4\kappa} > a \mid \frac{S_2-S_1}{\kappa} > d\right) = \exp(-4\kappa a)$$

because  $S_2-S_1-\kappa d \sim \exp(1)$  given the event  $\{S_2-S_1-\kappa d > 0\}$  (this is by the memoryless property of the exponential distribution).

This implies that given the event  $\{J_1=\mathcal{R}, J_2=Z\}$  the waiting time  $\hat{\tau}_3 \sim \exp(4\kappa)$  as in Barrio's SSAD and with probability 1  $J_3=\mathcal{R}$ . Continuing with Anderson-Kurtz's algorithm, at time  $\hat{t}_3$   $D$ ,  $Int$ ,  $S$  and  $x(\hat{t}_3)$  are updated to

$$D=\{d, \infty\}, \quad Int=4\kappa\hat{\tau}_1+\kappa d+4\kappa\hat{\tau}_3=S_2, \quad S=S_3 \quad x(\hat{t}_3)=1,$$

and  $\hat{\tau}_4$  is set as

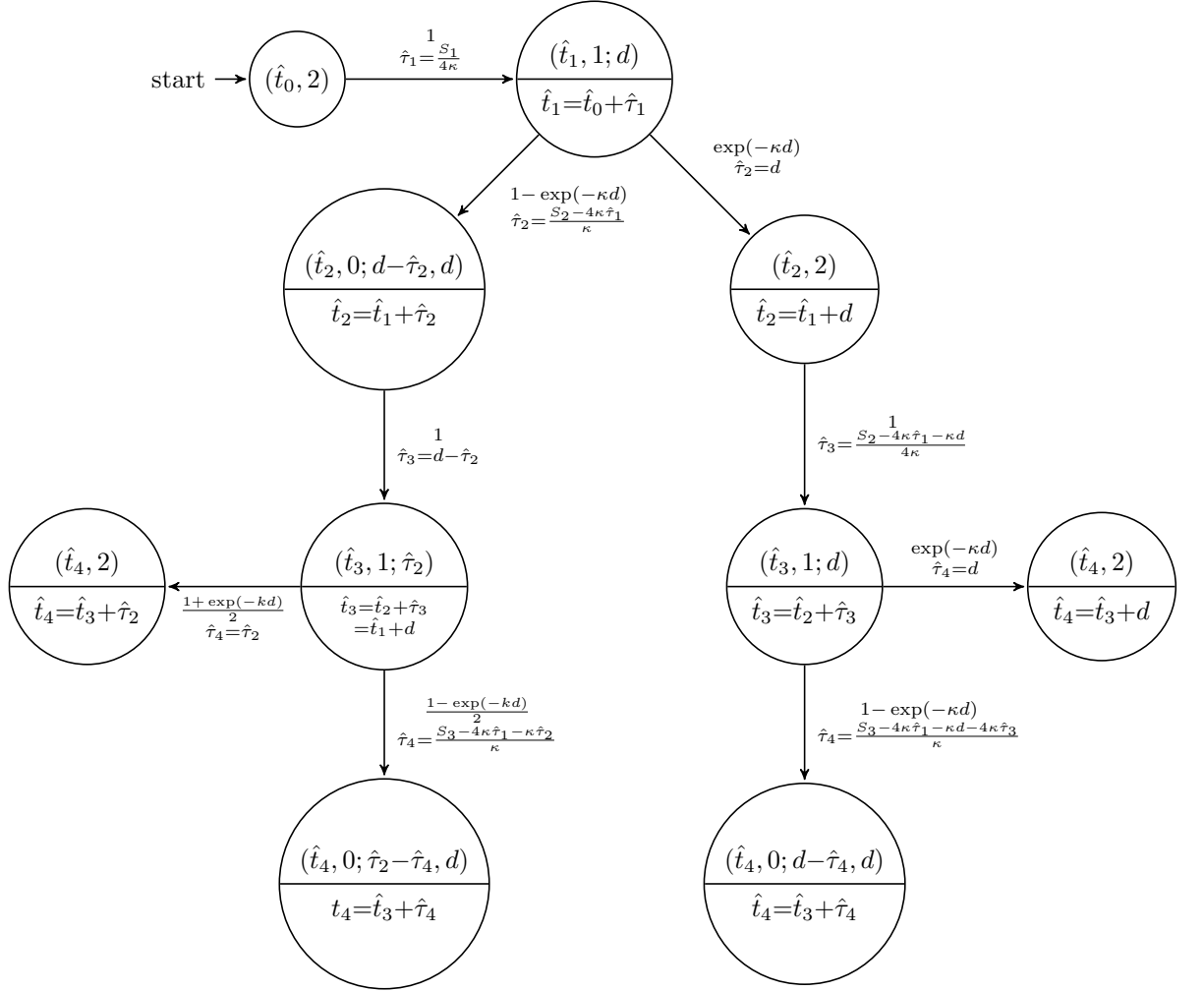
$$\hat{\tau}_4=\min\left(\left\{\frac{S-Int}{\kappa}\right\}\cup D\right)=\min\left(\frac{S_3-S_2}{\kappa}, d\right)$$

where  $\frac{S_3-S_2}{\kappa} \sim \exp(\kappa)$  and is independent of the past; that is, given the event  $\{J_1=\mathcal{R}, J_2=\mathcal{R}, J_3=\mathcal{R}\}$  the waiting time  $\hat{\tau}_4$  has the same probabilistic properties as in Barrio's SSAD.

The arguments above show that despite the difference between Barrio's SSAD and Anderson-Kurtz's SSAD, up to time  $\hat{t}_4$  the reaction jump process that Barrio's SSAD generates is a solution of the Anderson-Kurtz's equation

$$(5.9) \quad \begin{pmatrix} x(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^t \kappa(xw)(s) ds \right) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^{t-d} \kappa(xw)(s) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The following diagram shows the different paths that we have simulated above with Anderson-Kurtz's SSAD.

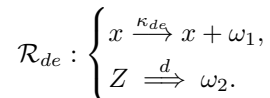


We use the same notation, the label  $\binom{p}{\hat{\tau}}$  on the arrows indicate the probability  $p$  of taking the particular transition and how the waiting time  $\hat{\tau}$  is calculated.

**Lemma 5.10.** *The paths calculated by Barrio's SSAD Algorithm are solutions to the Anderson-Kurtz's equation (5.9) up to time  $\hat{t}_4$ .*

5.4. **Cai's SSAD.** Another standard tool used in the simulation of chemical systems with delays is Cai's SSAD. Unlike Anderson-Kurtz's SSAD and Barrio's SSAD, Cai's SSAD only simulates those times where a chemical reaction will occur; see [5] for a historical presentation.

Consider again a chemical system of  $n$  chemical reactions  $\mathcal{R}_1, \dots, \mathcal{R}_n$  and one chemical reaction with delay  $\mathcal{R}_{de}$



Suppose that at time  $t$  the chemical system is in some state  $y$ , and that at this time there are  $l$  excited complexes  $Z$ , which will decompose at times

$$t+d_1 < t+d_2 \cdots < t+d_l.$$

If at time  $t+\tau$  occurs the next reaction, then Cai's proposal for the distribution of  $\tau$  is given by:  $\forall i \in \{0, \dots, l\}$ , and  $\forall s \in [d_i, d_{i+1})$

$$(5.11) \quad F_\tau(s) = 1 - \exp\left(-\sum_{j=0}^{i-1} \alpha_{sum,j}(d_{j+1}-d_j) - \alpha_{sum,i}(s-d_i)\right),$$

where  $d_0 := 0$ ,  $d_{l+1} := \infty$ , and  $\forall j \in \{0, \dots, l\}$

$$\alpha_{sum,j} := \sum_{\star \in \{1, \dots, n, de\}} \alpha_\star(y+j\omega_2).$$

As is shown in [5] formula (5.11) is motivated by Doob-Gillespie's SSA, and with this motivation in mind, Cai's proposal for the distribution of the type of reaction that occurs at time  $t+\tau$  is:  $\forall i \in \{0, \dots, l\}$

$$(5.12) \quad \mathbb{P}(J=\star | \tau \in [d_i, d_{i+1})) = \frac{\alpha_\star(y+i\omega_2)}{\alpha_{sum,i}}.$$

Note that if at time  $t$  there aren't excited complexes  $Z$  (that is  $l=0$ ), then  $\tau$  and  $J$  have the same distribution as in Doob-Gillespie's SSA. Also note that  $\forall 0 \leq s < d_1$ ,

$$\mathbb{P}(\tau \leq s) = 1 - \exp(-\alpha_{sum,0}s).$$

This means that in the time interval  $[0, d_1)$ , the waiting time  $\tau$  follows an exponential distribution as in Doob-Gillespie's SSA. Observe that this agrees with the fact that in the time interval  $[t, t+d_1)$  there are no excited complexes that will decompose, which implies that in this time interval the only possible stoichiometric change is a chemical reaction.

Another interesting point about the formula (5.11) is its sensitivity to stoichiometric changes due to decompositions of the excited complexes. For example observe how  $F_\tau$  passes from exponential distribution in the time interval  $[0, d_1)$ , to

$$F_\tau(s) = 1 - \exp(-\alpha_{sum,0}d_1 - \alpha_{sum,1}(s-d_1)) \quad \forall s \in [d_1, d_2)$$

due to the decomposition of an excited complex at time  $t+d_1$  (see [5]). Now we introduce Cai's SSAD; we refer to the reader to Pseudo-Code 1 in [5] for a method to simulate  $\tau$ . As in previous algorithms, all the information about the delayed reaction  $\mathcal{R}_{de}$  is in the index  $n+1$ .

**Algorithm 5.13.** (*Cai's SSAD*)

- (1) *Initialization.* Set  $y \leftarrow y_0$ ,  $t \leftarrow 0$  and  $D=\emptyset$ .
- (2) If  $D=\emptyset$ , then calculate  $\forall 1 \leq k \leq n+1$ ,  $\alpha_k(y)$  and set

$$\alpha_{sum,0} = \sum_{k=1}^{n+1} \alpha_k(y).$$

Otherwise, set  $l=|D|$  and calculate  $\forall 1 \leq k \leq n+1$ ,  $\forall 0 \leq i \leq l$ ,  $\alpha_k(y+l\omega_2)$ , and set

$$\alpha_{sum,i} = \sum_{k=1}^{n+1} \alpha_k(y+i\omega_2).$$

- (3) Generate  $\tau$  with distribution as (5.11).
- (4) Generate an uniform random variable  $u_1$ . If  $D=\emptyset$ , then set  $\mu$  to be the integer for which

$$\sum_{k=1}^{\mu-1} \frac{\alpha_j(y)}{\alpha_{sum,0}} \leq u_1 < \sum_{k=1}^{\mu} \frac{\alpha_k(y)}{\alpha_{sum,0}}.$$

Otherwise, if  $\tau \in [d_i, d_{i+1})$ , then set  $\mu$  to be the integer for which

$$\sum_{k=1}^{\mu-1} \frac{\alpha_j(y+i\omega_2)}{\alpha_{sum,i}} \leq u_1 < \sum_{k=1}^{\mu} \frac{\alpha_k(y+i\omega_2)}{\alpha_{sum,i}}.$$

If  $i \geq 1$ , then update  $y \leftarrow y+i\omega_2$  and update

$$D \leftarrow (D \setminus \{d_1, \dots, d_i\}) - \{\tau\}.$$

- (5) Update the state of the system  $y \leftarrow y+v[\mu]$ .
- (6) If  $\mu=n+1$ , then we update

$$D \leftarrow D \cup \{d\}.$$

Otherwise, go to the next step.

- (7) Set  $t \leftarrow t + \tau$ . Go to step 2.

From Cai's SSAD, the reaction jump process that this algorithm simulates has the following structure

$$(5.14) \quad Y(t) = \sum_{k=0}^{\infty} Y_k \chi_{[t_k, t_{k+1})}(t)$$

where the sequence  $\{t_k\}$  are the times where a chemical reaction takes place and the sequence  $\{Y_k\}$  satisfies  $\forall k \in \mathbb{N}_0$

$$Y_{k+1} = (Y_k + c(k)\omega_2) + w,$$

where  $w \in \{v[j], \omega_1 : 1 \leq j \leq n\}$  and  $c(k)$  is the number of excited complexes that decomposed in the interval  $[t_k, t_{k+1})$ .

Although the reaction jump process described above does not have the same structure that the Anderson-Kurtz SSAD, a question arises: Is the reaction jump process that Cai's SSAD generates a solution of the Anderson-Kurtz equation? This question is still open and in order to answer it we need much time and effort. In this section, we lay some foundations for further work. In particular, we must point out that trajectory (5.14) can be modified to include the effect of the physical transitions where the excited complexes  $Z$  decomposed. In this context, Cai asserted in [5] that his SSAD is statistically equivalent to Barrio's SSAD, but he does not give a complete proof of this fact. Our intention is to prove that the SSAD of Barrio and Cai are indeed equivalent; to do so we reinterpret Cai's SSAD inspired in Anderson-Kurtz's algorithm.

Recall that Doob-Gillespie's SSA has a clear interpretation as a competition between chemical reactions; see Theorem 4.22. This competition is expressed through a race of independent random variables with exponential distribution (see Chapter 4). Due to the structure of the distribution of the time between chemical reactions in Cai's SSAD (see (5.11)), a natural question arises: Can we endow Cai's SSAD with an interpretation in terms of a competition of chemical reactions? The answer to this question is affirmative as the following results show.

**Lemma 5.15.** *Let  $d_0 < d_1 < \dots < d_l < d_{l+1}$  be extended real numbers (they can be equal to  $\infty$ ) such that  $d_0 := 0$  and  $d_{l+1} := \infty$ . For each  $k \in \{1, \dots, n\}$ , let  $\alpha_{k,0}, \alpha_{k,1}, \dots, \alpha_{k,l}$  nonnegative real numbers such that  $\alpha_{k,l} > 0$*

$$\alpha_{sum,i} := \sum_{k=1}^n \alpha_{k,i} > 0 \quad \forall i \in \{0, \dots, l\}.$$

*Let  $X_1, \dots, X_n$  be independent random variables in a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , such that for each  $k \in \{1, \dots, n\}$ ,  $X_k$  has distribution*

$$\mathbb{P}(X_k \leq x) = \begin{cases} 1 - \exp(-\sum_{j=0}^{i-1} \alpha_{k,j}(d_{j+1} - d_j) - \alpha_{k,i}(x - d_i)) & x \in [d_i, d_{i+1}) \quad i \in \{0, \dots, l\} \\ 0 & x < 0. \end{cases}$$

*If  $Z = \min\{X_1, \dots, X_n\}$ , then  $Z$  has distribution*

$$\mathbb{P}(Z \leq x) = \begin{cases} 1 - \exp(-\sum_{j=0}^{i-1} \alpha_{sum,j}(d_{j+1} - d_j) - \alpha_{sum,i}(x - d_i)) & \text{if } x \in [d_i, d_{i+1}), \\ 0 & x < 0. \end{cases}$$

*Proof.* Let  $x \in [d_i, d_{i+1})$  for some  $i \in \{0, \dots, l\}$ . Then

$$\begin{aligned}
& \mathbb{P}(Z > x) \\
&= \mathbb{P}(X_1 > x, \dots, X_n > x) \\
&= \prod_{k=1}^n \mathbb{P}(X_k > x) \\
&= \prod_{k=1}^n \exp\left(-\sum_{j=0}^{i-1} \alpha_{k,j}(d_{j+1}-d_j) - \alpha_{k,i}(x-d_i)\right) \\
&= \exp\left(-\sum_{j=0}^{i-1} \alpha_{sum,j}(d_{j+1}-d_j) - \alpha_{sum,i}(x-d_i)\right).
\end{aligned}$$

□

**Theorem 5.16.** *Let  $X_1, \dots, X_n, Z$  be as in Lemma 5.15. If  $J$  is the index where the minimum is reached, that is if  $Z = X_J$ , then for those  $i \in \{0, \dots, l\}$  such that*

$$\mathbb{P}(d_i \leq Z < d_{i+1}) > 0,$$

*we get*

$$\forall r \in \{1, \dots, n\} \quad \mathbb{P}(J=r | d_i \leq Z < d_{i+1}) = \frac{\alpha_{k,i}}{\alpha_{sum,i}}.$$

*Proof.* Let  $r$  and  $i$  as above, and note that the following events are equivalent

$$\{J=r, d_i \leq Z < d_{i+1}\} = \{d_i \leq X_r < d_{i+1}, X_r \leq U\},$$

where

$$U = \min_{k \neq r} \{X_k\}.$$

Define  $\Psi : \mathbb{R} \times \Omega \rightarrow \mathbb{R}$

$$\Psi(x, \omega) = \begin{cases} 1 & d_i \leq x < d_{i+1} \quad x \leq U(\omega), \\ 0 & \text{in other case.} \end{cases}$$

Note that for every  $\omega \in \Omega$ , the function  $x \rightarrow \Psi(x, \omega)$  is piecewise continuous and for every  $x \in \mathbb{R}$ , the function  $\omega \rightarrow \Psi(x, \omega)$  is  $\sigma(U)$ -measurable. Then  $\Psi$  is  $\mathbb{B}(\mathbb{R}) \times \sigma(U)$ -measurable and observe that

$$E[\Psi(X_r, \cdot)] = \mathbb{P}(d_i \leq X_r < d_{i+1}, X_r \leq U)$$

Since  $X_r$  and  $U$  are independent, the freezing lemma 4.6 implies that

$$E[\Psi(X_r, \cdot) | X_r] = \Theta(X_r)$$

where  $\Theta(x) = E[\Psi(x, \cdot)]$ , that is

$$\Theta(x) = \begin{cases} \mathbb{P}(x \leq U) & \text{if } x \in [d_i, d_{i+1}), \\ 0 & \text{in other case.} \end{cases}$$

Therefore

$$\begin{aligned}
& \mathbb{P}(d_i \leq X_r < d_{i+1}, X_r \leq U) \\
&= E[\Psi(X_r, \cdot)] \\
&= E[E[\Psi(X_r, \cdot) | X_r]] \\
&= E[\Theta(X_r)] \\
&= \int_{d_i}^{d_{i+1}} \mathbb{P}(x \leq U) d(\mathbb{P} \circ X_r^{-1})(x).
\end{aligned}$$

From Lemma 5.15,

$$\forall x \in [d_i, d_{i+1}), \quad \mathbb{P}(x \leq U) = \exp \left( - \sum_{j=0}^{i-1} (\alpha_{sum,j} - \alpha_{r,j})(d_{j+1} - d_j) - (\alpha_{sum,i} - \alpha_{r,i})(x - d_i) \right)$$

and the density function of  $X_r$  satisfies

$$\forall x \in [d_i, d_{i+1}) \quad \alpha_{r,i} \exp \left( - \sum_{j=0}^{i-1} \alpha_{r,j}(d_{j+1} - d_j) - \alpha_{r,i}(x - d_i) \right).$$

Then

$$\begin{aligned}
& \mathbb{P}(d_i \leq X_r < d_{i+1}, X_r \leq U) \\
&= \int_{d_i}^{d_{i+1}} \mathbb{P}(x \leq U) d(\mathbb{P} \circ X_r^{-1})(x) \\
&= \int_{d_i}^{d_{i+1}} \mathbb{P}(x \leq U) \alpha_{r,i} \exp \left( - \sum_{j=0}^{i-1} \alpha_{r,j}(d_{j+1} - d_j) - \alpha_{r,i}(x - d_i) \right) dx \\
&= \int_{d_i}^{d_{i+1}} \alpha_{r,i} \exp \left( - \sum_{j=0}^{i-1} (\alpha_{sum,j})(d_{j+1} - d_j) - (\alpha_{sum,i} - \alpha_{r,i})(x - d_i) - \alpha_{r,i}(x - d_i) \right) dx \\
&= \alpha_{r,i} \exp \left( - \sum_{j=0}^{i-1} (\alpha_{sum,j})(d_{j+1} - d_j) \right) \int_{d_i}^{d_{i+1}} \exp(-\alpha_{sum,i}(x - d_i)) dx \\
&= \frac{\alpha_{r,i}}{\alpha_{sum,i}} \exp \left( - \sum_{j=0}^{i-1} (\alpha_{sum,j})(d_{j+1} - d_j) \right) (1 - \exp(-\alpha_{sum,i}(d_{i+1} - d_i))) \\
&= \frac{\alpha_{r,i}}{\alpha_{sum,i}} \left[ \exp \left( - \sum_{j=0}^{i-1} (\alpha_{sum,j})(d_{j+1} - d_j) \right) - \exp \left( - \sum_{j=0}^i (\alpha_{sum,j})(d_{j+1} - d_j) \right) \right],
\end{aligned}$$

where from lemma (5.15)

$$\mathbb{P}(d_i \leq Z < d_{i+1}) = \exp \left( - \sum_{j=0}^{i-1} (\alpha_{sum,j})(d_{j+1} - d_j) \right) - \exp \left( - \sum_{j=0}^i (\alpha_{sum,j})(d_{j+1} - d_j) \right).$$

It follows that

$$\mathbb{P}(J=r, d_i \leq Z < d_{i+1}) = \mathbb{P}(d_i \leq X_r < d_{i+1}, X_r \leq U) = \frac{\alpha_{r,i}}{\alpha_{sum,i}} \mathbb{P}(d_i \leq Z < d_{i+1}).$$

Therefore

$$\mathbb{P}(J=r | d_i \leq Z < d_{i+1}) = \frac{\alpha_{r,i}}{\alpha_{sum,i}}.$$

As we wanted to prove.  $\square$

The previous results are an advance in the direction to prove that Barrio and Cai's SSADs are statistically equivalent, in the sense that both algorithms produce trajectories of the form

$$(5.17) \quad \forall t > 0 \quad Y(t) = \sum_{j=0}^{\infty} Y_j \chi_{[\hat{t}_j, \hat{t}_{j+1})}(t)$$



where  $\{\hat{t}_j\}_j$  is the sequence of times when a stoichiometric change occurs and every  $Y_{j+1} := Y_j + w$  for  $w \in \{v[k], \omega_1, \omega_2 : 1 \leq k \leq n\}$ . Moreover, we assert that the equivalence between Barrio and Cai's SSADs happens, because the corresponding trajectories (5.17) are solutions to Anderson-Kurtz's equation

$$Y(t) = Y(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(Y(s)) ds \right) v_k \\ + \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \omega_1 + \mathcal{U}_{de} \left( \int_0^{t-d} \alpha_{de}(Y(s)) ds \right) \omega_2.$$

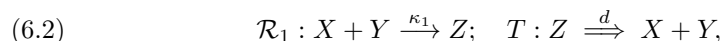
but this problem will be analyzed in the Doctoral studies.

## 6. CONCLUSIONS

The following results were developed and shown in this thesis work. We firstly presented in Section 2.1 and Chapter 3 a formal definition of a delayed chemical reaction, which was inspired by the theory of transition states in chemistry. Using this definition as a base, we then modify the original Anderson-Kurtz's equation for representing the stochastic time dynamics of a chemical reactor. For example, the following modified Anderson-Kurtz's equation

$$(6.1) \quad Y(t) = Y(0) + \sum_{k=1}^n \mathcal{U}_k \left( \int_0^t \alpha_k(Y(s)) ds \right) v[k] \\ + \mathcal{U}_{de} \left( \int_0^t \alpha_{de}(Y(s)) ds \right) \omega_1 + \mathcal{U}_{de} \left( \int_0^{t-d} \alpha_{de}(Y(s)) ds \right) \omega_2.$$

can model a system of  $n$  instantaneous chemical reactions and a delayed one. As a result we were able to show that the chemical system with one reaction  $\mathcal{R}_1$ , a physical transition  $T$  with delay  $d$



and initial state  $(1, y_0)^T$  with  $y_0 \in \mathbb{N}$ , can be modeled with a Semi-Markovian process which is the solution of the following Anderson-Kurtz's equation with delays

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 1 \\ y_0 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^t \kappa_1(xy)(s) ds \right) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^{t-d} \kappa_1(xy)(s) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The fact that we can find explicit solutions to a particular case of the delayed Anderson-Kurtz's equation (6.1) drives us to find more explicit solutions. Following the construction of the reaction jump process given in Chapter 2, in Section 4.1 we showed a new version of Doob-Gillespie's SSA, which has an interpretation in terms of a race between the chemical reactions of the system. We saw that this version of Doob-Gillespie's SSA is equivalent to Anderson-Kurtz's SSA, which allows us to show that for a chemical system with  $n$  chemical reactions, the corresponding reaction jump process is the unique solution to the stochastic equation (Anderson-Kurtz's equation)

$$X(t) = x_0 + \sum_{j=1}^n \mathcal{U}_j \left( \int_0^t \alpha_j(X(s)) ds \right) v[j].$$

The new interpretation of the Doob-Gillespie's SSA, as a race, gives a natural explanation on why the dynamics of a chemical reactor can be so efficiently represented by the original Doob-Gillespie's SSA. Moreover, the equivalence between Doob-Gillespie's SSA and Anderson-Kurtz's SSA implies that Anderson-Kurtz's SSAD (which generates the solution to equation (6.1)) can be interpreted as a modification to the Doob-Gillespie's SSA. We also noted in Chapter 5 that Barrio's SSAD and Cai's SSAD are modifications to Doob-Gillespie's SSA. Thus, a natural question arises: For a chemical system of  $n$  instantaneous chemical reactions and a delayed one, can the trajectories of the solution to equation (6.1) be simulated through

Barrio's SSAD and Cai's SSAD? We believe that the answer to this question is affirmative, and although we could not give a proof for this statement, in Section 5.3 we saw that for the system (6.2) with initial conditions  $(2, 2)^T$ , Barrio's SSAD generates the trajectories, until time  $\hat{t}_4$ , of the solution to the Anderson-Kurtz's equation with delays

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^t \kappa_1(xy)(s) ds \right) \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mathcal{U}_1 \left( \int_0^{t-d} \kappa_1(xy)(s) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Finally, in order to lay the foundations for future work, we prove in Section 5.4 that Cai's SSAD also has a natural interpretation in terms of a race between the different chemical reactions participating in the system. We assert that this interpretation is a forward step into proving that the trajectories generated by Cai's SSAD are solutions to the modified Anderson-Kurtz's equation (6.1). These ideas deserve to be further developed in a more detailed Doctoral work.

## REFERENCES

- [1] D. F. ANDERSON, *A modified next reaction method for simulating chemical systems with time dependent propensities and delays*, The Journal of Chemical Physics, 127 (2007), p. 214107.
- [2] P. BALDI, *Stochastic Calculus: An Introduction Through Theory and Exercises*, Springer, 1 ed., 2017.
- [3] M. BARRIO, K. BURRAGE, A. LEIER, AND T. TIAN, *Oscillatory regulation of hes1: Discrete stochastic delay modelling and simulation*, PLoS Comput. Biol., 2 (2006), p. e117.
- [4] P. BRÉMAUD, *Markov Chains, Gibbs Fields, Monte Carlo Simulation and Queues*, Springer, 1 ed., 2001.
- [5] X. CAI, *Exact stochastic simulation of coupled chemical reactions with delays*, J. Chem. Phys., 126 (2007), p. 124108.
- [6] J. L. DOOB, *Topics in the theory of markoff chains*, Trans. Am. Math. Soc., 52 (1942), pp. 37–64.
- [7] ———, *Markoff chains – denumerable case*, Trans. Am. Math. Soc., 58 (1945), pp. 455–473.
- [8] D. F. ANDERSON AND T. G. KURTZ, *Stochastic Analysis of Biochemical Systems*, Springer, 1 ed., 2015.
- [9] D. T. GILLESPIE, *The chemical langevin equation*, J. Phys. Chem., 113 (2000), pp. 297–306.
- [10] K. GOWRISANKARAN, *Measurability of functions in product spaces*, AMERICAN MATHEMATICAL SOCIETY, 31 (1972), pp. 485–488.
- [11] S. A. HARICH, D. DAI, C. C. WANG, X. YANG, S. D. CHAO, AND R. T. SKODJE, *Forward scattering due to slow-down of the intermediate in the  $H + HD \rightarrow D + H_2$  reaction*, Nature, 419 (2002), pp. 281–284.
- [12] D. J. HIGHAM, *Modeling and simulating chemical reactions*, SIAM REVIEW, 50 (2008), pp. 347–368.
- [13] T. G. KURTZ, *Solutions of ordinary differential equations as limits of pure jump markov processes*, Journal of Applied Probability, 7 (1970), pp. 49–58.
- [14] ———, *Limit theorems for sequences of jump markov processes approximating ordinary differential processes*, Journal of Applied Probability, 8 (1971), pp. 344–356.
- [15] ———, *Limit theorems and diffusion approximations for density dependent markov chains*, Mathematical Programming Study North-Holland Publishing Company, 5 (1976), pp. 67–78.
- [16] ———, *Strong approximation theorems for density dependent markov chains*, Stochastic Processes and their Applications North-Holland Publishing Company, 6 (1978), pp. 223–240.

- [17] ———, *Representations of markov processes as multiparameter time changes*, The Annals of Probability, 8 (1980), pp. 682–715.
- [18] D. E. MANOLOPOULOS, *A delayed reaction*, Nature, 419 (2002), pp. 266–267.
- [19] S. M. ZERON ES, *Numerical solution of the chemical master equation uniqueness and stability of the stationary distribution for chemical networks, and mrna bursting in a gene network with negative feedback regulation*. *methods enzymol*, (2001).