Stochastic Models of Gene Expression

by

Lanjia Lin

Bachelor of Engineering
Wuhan University, 2003

Submitted in Partial Fulfillment of the

Requirements for the Degree of Master of Science

in the Department of Mathematics

University of South Carolina

2005

_____   _____
Department of Mathematics   Department of Mathematics
Director of Thesis   2nd Reader

_____
Dean of the Graduate School

Acknowledgements

First, I would like to express my gratitude to my advisor, Dr. Daniel Dix, for his invaluable guidance and constant patience. Second, I wish to express my thanks to Dr. Laszlo Szekely for reading my thesis and offering helpful suggestions. Also, I would like to thank the faculty and staff in the Department of Mathematics for their help and support during my study here.

Finally, I would like to thank my wonderful parents, Shaokang and Yonglan, for their selfless love and continuous support.

Abstract

This thesis addresses the problem of describing the time evolution of a genetic chemical system using a stochastic approach: the Master equation approach, which is the consequence of the fundamental hypothesis of the stochastic formulation of chemical kinetics. The Master equation can be used to compute the time dependent probability distribution of the state of a chemical system. The Master equation is built on a Markov model foundation. We derive the Master equation under the assumption that there exits a Markov process satisfying three particular postulates.

A mathematical formulism called Petri net which represents a chemical system is introduced. The Petri net and a positive integer k together determine a State graph. We also explore the properties of the solution of the Master equation associated with the State graph.

We give a construction of a Markov process on an infinite product space and prove that the constructed Markov process satisfies the three postulates, hence satisfies the Master equation.

We also give the time evolution outcomes of the stochastic simulations of a genetic auto-regulatory system using computer software, which are consistent with the long time behavior property of the solution of the Master equation.

Contents

List of Figures

Chapter 1

Introduction

The process linking genes to proteins is called **Gene Expression**. The details of the biology background of processes of gene expression are given in Chapter 2. In a genetic regulatory system, it is possible for the protein produced from a gene to influence the rate of expression of that same gene. A genetic regulatory system can be represented by the chemical species and reactions involved in the regulatory process, hence can be considered as a homogeneous chemical system.

In Chapter 3 we introduce a mathematical representation for chemical systems called **Petri net**, and discuss how to describe the time behavior of a chemical system in a mathematical way. A Petri net for a chemical system determines an oriented bipartite graph whose edges are labeled by numbers C specifying the reaction rates; A Petri net and an integer $k > 0$ together determine a directed graph called the **State graph** whose oriented edges are labeled by numbers W specifying the probabilities of the occurrences of reactions per unit time. It is shown that a system of ordinary differential equations called the **Reaction Rate equation** is completely determined by the labeled Petri net. The relationship between functions C and W is given in Section 3.2. There are two mathematical approaches to describe the time evolution of a homogeneous chemical system: the deterministic approach via the Reation Rate equations and the stochastic approach. The **fundamental hypothesis of the stochastic formulation of chemical kinetics** and relationship between the deterministic ap-

proach and stochastic approach are discussed in Section 3.2. The stochastic nature of a genetic network in single cells cannot be neglected when modeling cellular control circuits. Hence a **Markov process** is needed to describe the stochastic trajectories of the system behavior.

In Chapter 4, a system of differential equations called the **Master Equations** is derived, which can be used to describe the stochastic model of a chemical system and to compute the time dependent probability distribution of the state of the system. It is shown that under the assumption of the existence of a Markov process whose conditional probabilities satisfy three particular postulates, the probability distribution of the Markov process satisfies the Master equation. The properties of the solution of the Master equation associated with an abstract directed graph are given and the long time behavior of the solution of the Master equation is explored.

We construct the Markov process which can describe the system behavior in Chapter 5. Also the three postulates of the constructed process have been verified, which complete the proof that the time dependent probability distribution of the state of the system satisfies the Master equation.
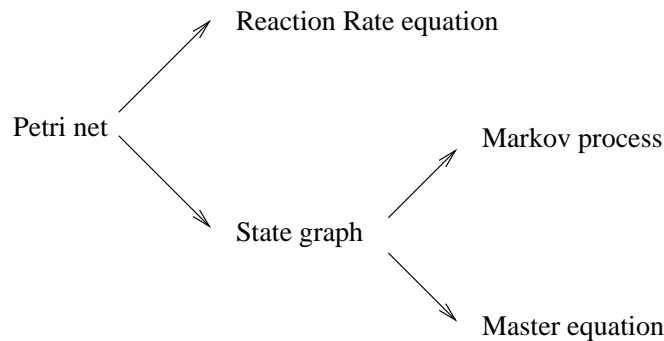
Figure 1.1: Relationships between Petri Net, Reaction Rate Equation, etc.

Figure 1.1 shows the relationships between Petri net, Reaction Rate equation, State graph, Markov process and Master equation: A Petri net determines a Reaction

Rate equation and a State graph; a State graph determines a Markov process and a Master equation; the relationship between the Reaction Rate equation and the Master equation is discussed in Section 3.2.

A model of a auto-regulating genetic system is constructed and studied in Chapter 6. The corresponding Master equation is then written down.

In Chapter 7 we simulate the time evolution of a simplified model of an auto-regulating genetic system using a biochemical simulation software called *Jarnac*, which uses the Gillespie algorithm. Several outcomes of the simulation have been given and discussed. The behavior observed in the simulation coincides with that predicted in the theorem of the long time behavior of the solution of the Master equation we give in Chapter 4.

Chapter 2

Overview of Processes Linking Genes to Proteins

All prokaryotes and eukaryotes are composed of cells, and the inner workings of all cells are remarkably similar, being dominated by the production and reactions of biopolymers: nucleic acids, proteins, and carbohydrates. A polymer is a long chainlike molecule consisting of many identical or similar building blocks linked by covalent bonds, much as a train consists of a chain of cars.

Proteins are the most structurally sophisticated molecules known and the molecular tools of the cell. Consistent with their diverse functions, they vary extensively in structure, each type of protein having a unique three-dimensional shape. Diverse though proteins may be, they are all polymers constructed from the same set of 20 amino acids. Amino acids are organic molecules possessing both carboxyl and amino groups. Polymers of amino acids are called polypeptides. A protein consists of one or more polypeptide folded and coiled into specific conformations. More details of the protein structure can be found in ([2]).

In this chapter, we will overview the processes of production of protein from the corresponding gene to provide the biological setting for the type of mathematical models we will study. Genes provide the instructions for making specific proteins. Nucleic acids and proteins have specific sequences of monomers that convey informa-

tion. In DNA or RNA, the monomers are the four types of nucleotides, which differ in their nitrogenous bases. Each gene has a specific sequence of bases, which might be hundreds or thousands of nucleotides long. The monomers of proteins are the 20 amino acids. Thus nucleic acids and proteins contain information written in two different chemical languages. A cell does not directly translate a gene into a chain of amino acids. The bridge between DNA and protein synthesis is RNA.

To get from DNA to protein requires two major stages, transcription, and translation. In this thesis, we only discuss how these two processes take place in prokaryotes, which is much easier than the situation in eukaryotes. The following details are adapted from ([4]).

Transcription

Transcription is the synthesis of RNA under the direction of DNA. Both types of nucleic acids use the same language, except thymine in DNA is converted to uracil in RNA, and deoxyribonuclieotides are replaced by ribonucleotides. The information is simply transcribed from one molecule to the other. A DNA strand provides a template for assembling a sequence of RNA nucleotides. An enzyme called an **RNA polymerase (RNAP)** pries the two strands of DNA apart and hooks together the RNA nucleotides as they pair along the DNA template. The type of RNA molecule created in this manner is called **messenger RNA (mRNA)**. The three stages of transcription are initiation, elongation, and termination of the RNA chain, as shown in Figure 2.1.

Initiation: At the first stage, RNA polymerase binds to some region of DNA. The region where RNA polymerase attaches and initiates transcription is called a **promoter**. In addition to determining where transcription starts, the promoter also determines which of the two strands of the DNA helix is used as the template, and therefore the direction along duplex DNA in which transcription will proceed. The
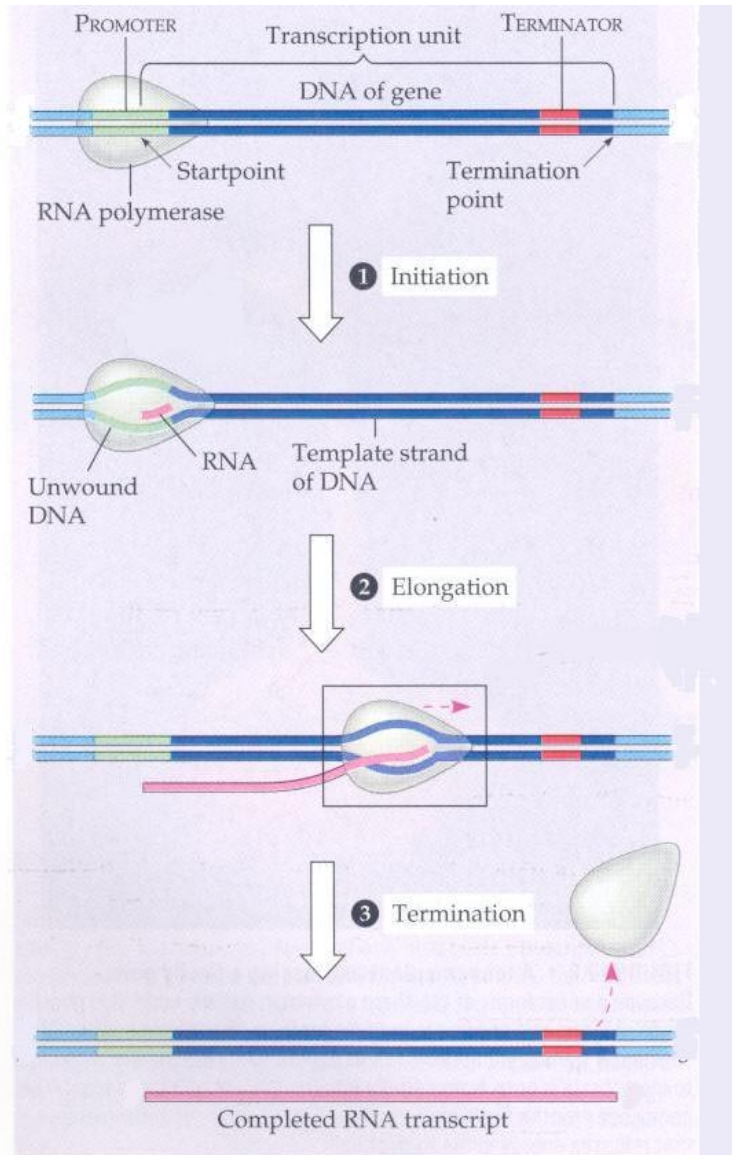
Figure 2.1: Transcription Processes

completed assembly of promoter and RNA polymerase is called a transcription initiation complex.

Elongation: As RNA polymerase moves along the DNA, it continues to untwist the double helix. One of the strands provides a template for assembling a sequence of RNA nucleotides. The enzyme adds a nucleotide to only one end of the growing RNA molecule as it moves along the double helix. In the wake of this advancing wave of RNA synthesis, the DNA double helix reforms and the new RNA molecule peels away from its DNA template. A single gene can be transcribed successively by several molecules of RNAP following each other. The growing strands of RNA trail off from each polymerase, with the length of each new strand reflecting how far along the template the enzyme has traveled from the start-point.

Termination: Transcription proceeds until shortly after the RNA polymerase transcribes a DNA sequence called a **Terminator**. The transcribed terminator, which is an RNA sequence, functions as the actual termination signal, which leads to dissociation of RNAP from DNA and complete reformation of the double helix of DNA.

In the genetic code, the genetic instructions for a polypeptide chain are written in the DNA as a series of three-nucleotide words. During transcription, the gene determines the sequence of base triplets along the length of an mRNA molecule. The mRNA base triplets are called **Codons**. Each codon specifies which one of the 20 amino acids will be incorporated at the corresponding position along a polypeptide by translation.

An mRNA's stability is controlled by the RNA degradation enzyme, for example, the one which is called RNase E in phage $\lambda-$infected *Escherichia coli* cells. The RNA degradation enzyme can initiate degradation (i.e. chemical decay) of the mRNA when it binds to mRNA.

The agents of translation of mRNA are large molecules called **Ribosomes**. RNA degradation enzyme cannot bind to mRNA when mRNA is bound by a Ribosome, which can prevent the mRNA from degradating until the site is again exposed as the Ribosome finishes translating the mRNA. At each exposure of the Ribosome binding site and RNA degradation enzyme binding site on mRNA, there is a direct competition between Ribosome and RNA degradation enzyme binding. This competition leads either to successful production of a protein or to degradation of the mRNA.

Translation

Translation is the actual synthesis of a polypeptide under the direction of an mRNA. The cell interprets a genetic message and builds a protein accordingly. The interpreter is called **transfer RNA (tRNA)**. The function of tRNA is to transfer amino acids from the cytoplasm's amino acid pool to a ribosome. The Ribosome then adds each amino acid brought to it by tRNA to the growing end of a polypeptide chain. The key to translating a genetic message into a specific amino acid sequence is that each type of tRNA molecule links a particular mRNA codon with a particular amino acid.

Ribosomes facilitate the specific coupling of tRNA anticodons with mRNA codons during protein synthesis. A Ribosome is made up of two subunits, termed the large and small subunits. The structure of a Ribosome reflects its function of bringing mRNA together with two amino acids, as shown in Figure 2.2. In addition to a binding site for mRNA, each Ribosome has three binding sites for tRNA. The P site holds the tRNA carrying the growing polypeptide chain, while the A site holds the tRNA carrying the next amino acid to be added to the chain. Discharged tRNAs leave the Ribosome from the E site.

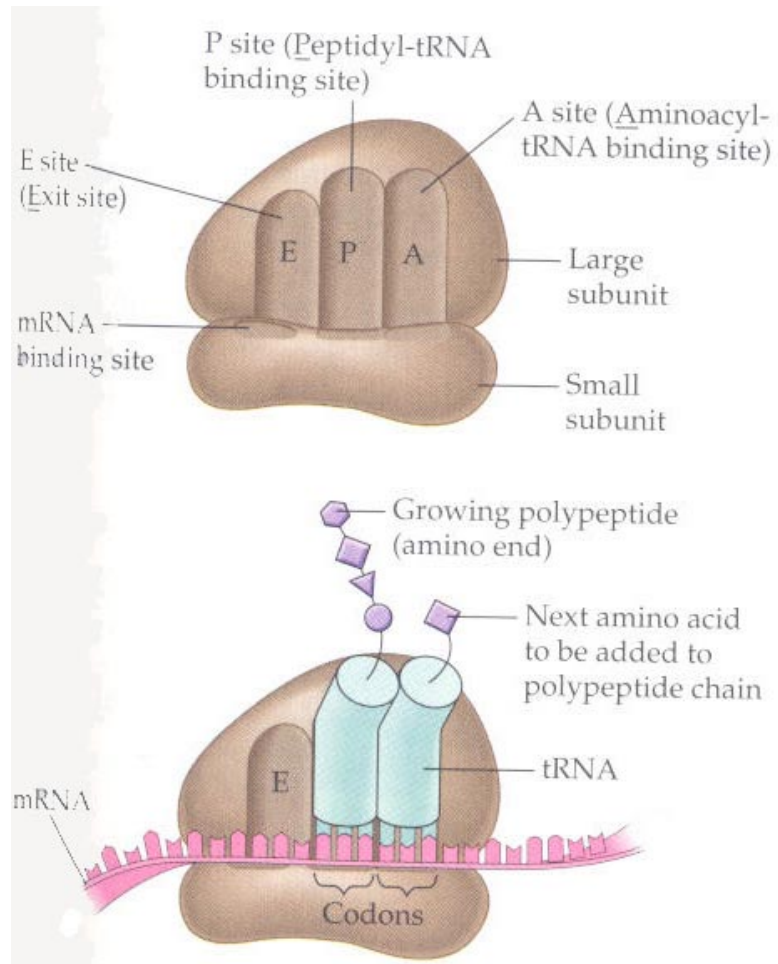Like transcription, translation can be divided into three stages: initiation, elongation and termination.

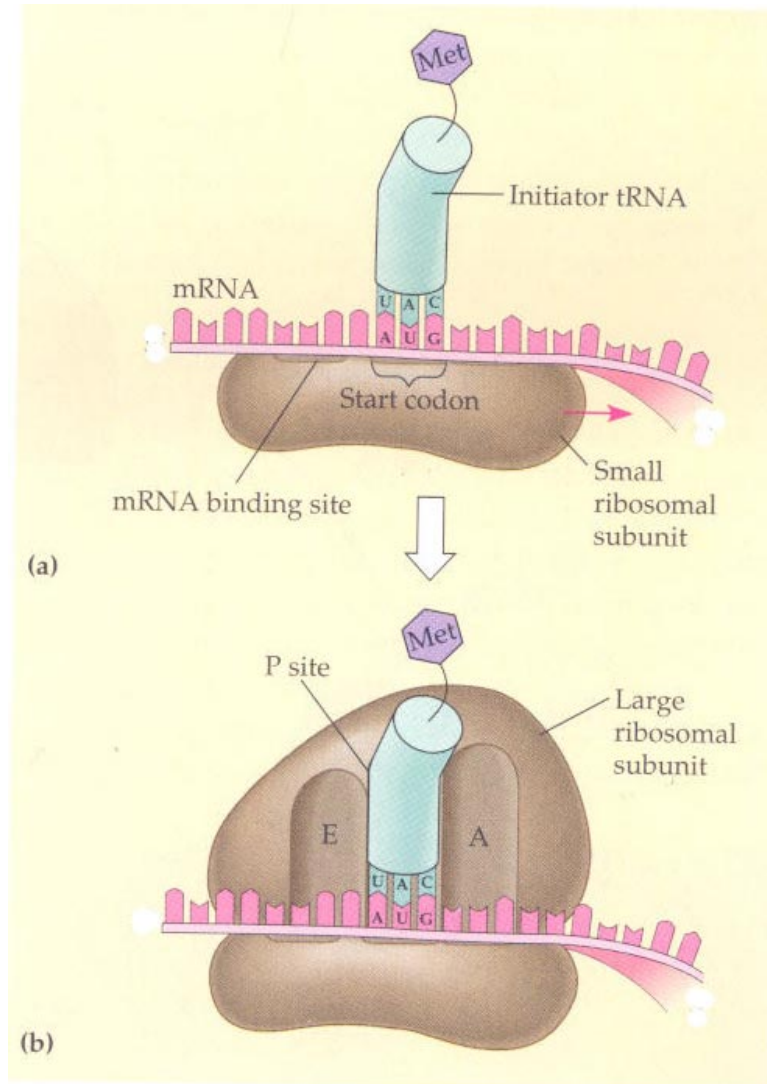Figure 2.2: Structure of Ribosome

Figure 2.3: Initiation Process of Translation
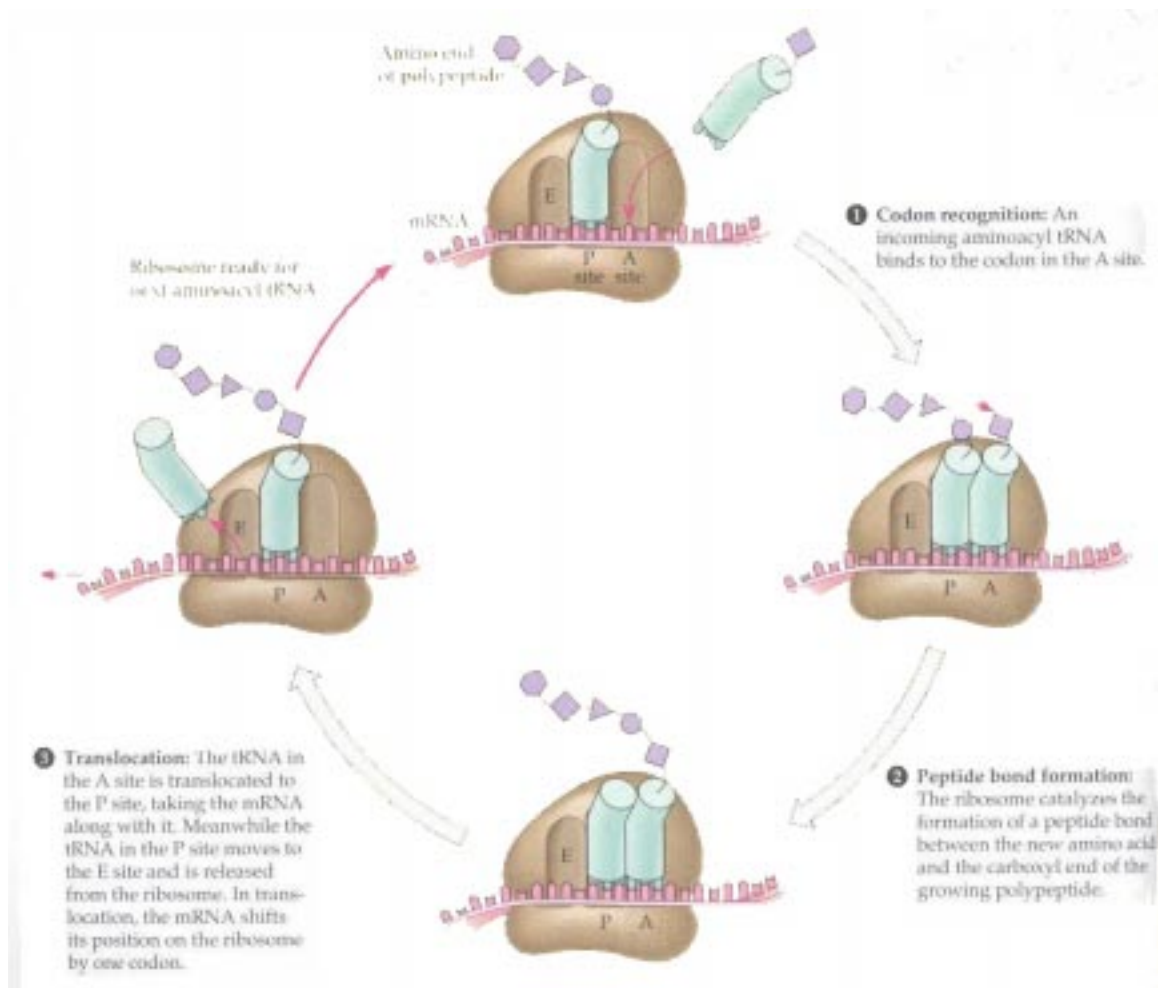
Figure 2.4: Elongation Process of Translation

Initiation: This stage brings together mRNA, a tRNA bearing the first amino acid of the polypeptide, and the two subunits of a Ribosome, as shown in Figure 2.3. First, a small Ribosome subunit binds to both the leader segment of the mRNA and a special initiator tRNA. Then it is followed by the attachment of a large Ribosomal subunit, completing a translation initiation complex.

Elongation: Amino acids are added one by one to the first amino acid in this stage, as shown in Figure 2.4. Each addition occurs in a three-step cycle: codon recognition, peptide bond formation, and translocation. First the mRNA codon in the A site of the Ribosome forms hydrogen bonds with the incoming tRNA carrying its appropriate amino acid. Then the large Ribosome subunit catalyzes the formation of a peptide bond that joins the polypeptide extending from the P site to the newly arrived acid in the A site. At last, the tRNA in the A site is translocated to the P site. The mRNA moves along with it and brings the next codon to be translated into the A site. Meanwhile, the tRNA that was in the P site moves to the E site, disconnects from mRNA and leaves the Ribosome.

Termination: Elongation continues until a *stop codon* reaches the A site of the Ribosome. The polypeptide is released from the Ribosome. The remainder of the translation assembly then comes apart.

It is possible for the protein produced from a gene to influence the rate of transcription of that same gene. Such a genetic system is called the *auto-regulating genetic system*, which will be discussed in Chapter 6. A genetic regulatory system can be considered as a chemical system.

Chapter 3

Stochastic Basis of Chemical Models

3.1 Representation of Molecular Interactions as Petri Nets

A system of molecular interactions can be represented as a mathematical formalism called a labeled **Petri net**. A Petri net (S, R, A, B) is defined as follows:

(1) S and R are finite sets with $S \cap R = \emptyset$;

(2) $A \subset S \times R$ and $B \subset R \times S$.

where elements of S are called *chemical species*, elements of R are called *chemical reactions*, $A \cup B$ is called the *flow relation* ([17]).

A labeling (I, O, C) of the Petri net (S, R, A, B) consists of functions:

(1) $I : A \to \mathbb{N} \cup \{0\}$ and $O : B \to \mathbb{N} \cup \{0\}$;

(2) $C : R \to (0, \infty)$.

where functions I and O are called *input* and *output functions*, and the function C specifies the reaction rate constants ([11]).

A Petri net determines an oriented bipartite graph $(S \cup R, A \cup B)$, where species are drawn as circles, reactions as rounded rectangles, and members of $A \cup B$ are drawn

as arrows, referred to as directed arcs (see Fig 3.1). Values of the input and output functions greater than one are indicated as labels of the corresponding directed arcs ([11]).

To express a chemical system as a Petri net, let A be the set of all pairs (s, r) where s is a chemical species required as a reactant in the chemical reaction r, and B be the set of all pairs (r, s) where s is a chemical species produced as a result of the chemical reaction r. The input and output functions capture the stoichiometry of all the reactions being considered, i.e., if (s, r) is in A then $I(s, r)$ gives the number of molecules of reactant s consumed through reaction r, and for (r, s) in B, $O(r, s)$ denotes the number of molecules of product s produced through reaction r. We also assume for every reaction $r \in R$ that the integers

$$\{I(s, r) \mid (s, r) \in A\} \cup \{O(r, s) \mid (r, s) \in B\}$$

have no common prime factors.

Let us take the following reaction as an example:

$$r : 2s_1 + s_2 \longrightarrow 2s_3.$$
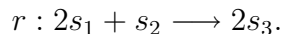
We have $I(s_1, r) = 2$, $I(s_2, r) = 1$, and $O(r, s_3) = 2$. These three integers have no common prime factors. We could also represent the reaction r as:

$$r : 4s_1 + 2s_2 \longrightarrow 4s_3.$$
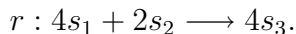
For simplicity, however, we will not do so.

A *state* of the Petri net (S, R, A, B) is a mapping $\mathcal{X} : S \longrightarrow \mathbb{N} \cup \{0\}$. $\mathcal{X}(s)$ specifies the number of molecules of species s present in the system. After a reaction occurs, the molecule numbers of the input species decrease and molecule numbers of the output species increase according to the stoichiometric coefficients. Let $\mathbf{n}_r \in \mathbb{Z}^S$

denote the difference between the states of the system before and after reaction r, then

$$\mathbf{n}_r(s) = \begin{cases} I(s,r) & \text{if } (s,r) \in A, \\ -O(r,s) & \text{if } (r,s) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Hence if state $\mathcal{X}$ is reached from state $\mathcal{X}'$ after the occurrence of reaction r, we have $\mathcal{X}' = \mathcal{X} + \mathbf{n}_r$. We assume that the mapping $R \longrightarrow \mathbb{Z}^S : r \longmapsto \mathbf{n}_r$ is one to one. That is, if two reactions have the same values for the vector $\mathbf{n}_r$, then we simply consider these two reactions to be identical.

The following is an example of a Petri net for the dimerization reactions $2P \rightleftharpoons P_2$. Let s_1 denote P, s_2 denote P_2, r_1 represents $2P \rightarrow P_2$, and r_2 represents $P_2 \rightarrow 2P$, we have:

$$S = \{s_1, s_2\};$$
$$R = \{r_1, r_2\};$$
$$A = \{(s_1, r_1), (s_2, r_2)\};$$
$$B = \{(r_1, s_2), (r_2, s_1)\};$$
$$I(s_1, r_1) = 2, I(s_2, r_2) = 1, O(r_1, s_2) = 1, O(r_2, s_1) = 2.$$

And $\mathbf{n}_{r_1} = (2, -1), \mathbf{n}_{r_2} = (-2, 1)$, where $\mathbf{n}_r = (\mathbf{n}_r(s_1), \mathbf{n}_r(s_2))$.

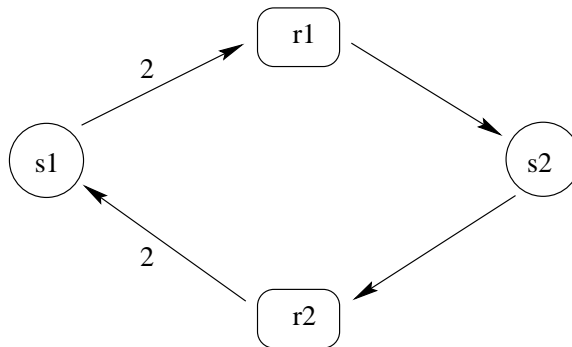The corresponding graphical form of the Petri net is shown as Fig 3.1:

Figure 3.1: Petri Net for Dimerization Reactions $2P \rightleftharpoons P_2$.

A Petri net (S, R, A, B) and an integer $k > 0$ together determine a directed graph $G = (\mathcal{V}, E)$ with no loops as follows. Let k represent the maximum value of the number of molecules of any species allowed, and $\mathcal{V} = \{0, 1, ..., k\}^S$ be the finite set of vertices of G, each element of which is a state of the system. Define $\Delta = \{(v, v) \mid v \in \mathcal{V}\} \subset \mathcal{V} \times \mathcal{V}$, and let $E \in (\mathcal{V} \times \mathcal{V}) \backslash \Delta$ be the following set of oriented edges:

$$E = \{(\mathcal{X}', \mathcal{X}) \in \mathcal{V} \times \mathcal{V} \mid \exists r \in R, \mathcal{X}' = \mathcal{X} + \mathbf{n}_r\}.$$

Hence a reaction r might determine an oriented edge $(\mathcal{X}', \mathcal{X})$ connecting the two vertices $\mathcal{X}' = \mathcal{X} + \mathbf{n}_r$ and $\mathcal{X}$. $G = (\mathcal{V}, E)$ is called the State graph associated to the Petri net (S, R, A, B) and the integer k.

3.2 Fundamental Hypothesis of the Stochastic Formulation of Chemical Kinetics

There are two mathematical models to describe the time evolution of a homogeneous chemical system: the deterministic model and the stochastic model. The following discussion of the differences between these two models is adapted from ([9]). The deterministic model assumes the change of concentration of each chemical species over time is a continuous and deterministic process, governed by a set of coupled ordinary differential equations, which are called the **Reaction-Rate Equations**. In this model, the system will always evolve in an identical way if given the same initial concentration vector. By the contrast, the stochastic model regards the time evolution of the system as a discrete and stochastic process, where two independent simulations of the same system with the same initial state would generate different time evolutions. The stochastic model is formulated in terms of probabilities, which are governed by a system of ordinary differential equations, the **Master Equation**.

The simplest way to view a stochastic model is to take it as a device for ignoring

all the physical factors and details which are necessary to make a deterministic pre-diction of the future. To illustrate what we mean, suppose using all the physical laws governing molecular motions and interactions we construct a 'super model' which ac-counts for the positions and velocities of all the atoms in a large system with a fixed volume V. Assume equations of motion are formulated, whose solutions give the time dependent trajectories of the system ([6]) ([10]).

Suppose there are N particles in the system. Let $\Phi = \mathbb{R}^{6N}$ be the phase space of this system; each point ϕ in Φ specifies all the positions and the velocities of all the particles in the system at some instant of time. Let the function H be defined on Φ giving the total energy of the system; here we assume that the system is isolated. If given a specified ϕ of the system at time 0, then the trajectory function $\psi_t(\phi)$ is uniquely determined for all time, and $H(\psi_t(\phi))$ is independent of t.

Let

$$Z(\theta) = \int_\Phi \exp[-H(\phi)/(k_B\theta)]\,d\phi$$

where k_B is Boltzmann's constant, θ is the absolute temperature of the system, and $d\phi$ is the Lebesgue measure on Φ. Then

$$\rho(\phi) = \exp[-H(\phi)/(k_B\theta)]/Z(\theta)$$

defines a probability density function on Φ such that the probability of a subset A of Φ coincides with the probability of the subset $\psi_t(A)$ for all t.

Suppose using the particles within the system it is possible to form various chem-ical species s in S, where S is the set of all chemical species to be considered. Let $\mathcal{X}_s$ be a function on Φ which assigns to each ϕ in Φ the number of copies of the molec-ular species s present in the system as described by ϕ. Thus $\mathcal{X}_s(\psi_t(\phi))$ is a function which describes how this number of copies changes as time evolves. $\mathcal{X}_s(\psi_t(\phi))$ is a stochastic process on the probability space $(\Phi, \rho\,d\phi)$, but it is not Markovian, see

section 3.3. Since $\mathcal{X}_s$ has values in the nonnegative integers this function will have discontinuities. Note that it is impossible for us to know the initial ϕ. What we know about ϕ is the above probability function on Φ. Each random sample ϕ from this probability distribution generates a distinct time history of the number of molecules of the species.

In the following we will adopt the notation we used in Petri net representation. Suppose various chemical reactions r from set R are possible in our system. A chemical reaction r must take place between actual molecules of species s where (s, r) is in A; $I(s, r)$ actual molecules are required for each species s involved. There are $\binom{\mathcal{X}_s(\phi)}{I(s,r)}$ different ways to group the $\mathcal{X}_s(\phi)$ molecules of species s into groups of size $I(s, r)$. Thus there are

$$\prod_{(s,r)\in A} \binom{\mathcal{X}_s(\phi)}{I(s, r)}$$

different ways the reaction r could potentially happen in the system as described by ϕ in Φ.

However in many of these groupings the molecules are not sufficiently close together in space for the reaction to occur. Suppose all the reactant molecules need to be in a volume of size V_r in order for the reaction r to proceed, and the volume V is divided into V/V_r cells each of volume V_r. There are

$$l_r = \sum_{(s,r)\in A} I(s, r)$$

reactant molecules taking part in reaction r. If these are randomly assigned to cells, the first assignment determines the common cell, and for each of the $l_r - 1$ remaining assignments the fraction V_r/V land in the common cell. Hence the probability that a given combination of l_r reactant molecules will end up close enough together to react is proportional to $(V_r/V)^{l_r-1}$. However the quantity V_r is difficult to know exactly.

Even if the reactant molecules are close enough together in space they might not be travelling in the correct directions so that the reaction can proceed. Even if they are close enough together in space and travelling in the right directions they may not be moving fast enough to overcome some potential energy barriers so that the reaction can proceed. This factor is influenced by the temperature θ, which is fixed in our discussion. Thus it is very difficult to quantify the necessary physical conditions to calculate which of the possible groupings will actually lead to the reaction r when the system is described by ϕ. However we will assume that it is possible to count the number of times the chemical reaction r actually occurs during a time interval $(t, t+dt)$ for a particular system trajectory $\psi_t(\phi)$; let this number be denoted by $N(r, t, dt, \phi)$. Then the expected number of times the reaction r occurs per theoretical opportunity in the time interval $(t, t + dt)$ is

$$\int_\Phi \frac{N(r, t, dt, \phi)}{\displaystyle\prod_{(s,r)\in A} \binom{\mathcal{X}_s(\psi_t(\phi))}{I(s,r)}} \rho(\phi) \, d\phi.$$

This quantity should be roughly proportional to $V^{1-l_r} dt$ as long as dt is small but not as small as the time required for the reaction r to occur. The **Fundamental Hypothesis of the Stochastic Formulation of Chemical Kinetics** is that there is a constant $C(r)$ (independent of time t) such that the above quantity is $C(r)V^{1-l_r} dt + o(dt)$ as dt tends to zero ([8]) ([15]). The constant $C(r)$ will in general depend upon the temperature θ but will not depend on the volume V. Obviously this hypothesis cannot be true when dt is as small as the time required for the reaction r to occur, but we take this as an assumption giving rise to a model whose usefulness must be tested empirically. We may understand $C(r)V^{1-l_r} dt$, for dt small, as the

following:

$$C(r)V^{1-l_r}dt = \text{to the first order in } dt, \text{ the probability that a particular combination}$$

of the reactant molecules will undergo reaction r in any time

interval of duration dt. (3.1)

where $C(r)$ contains in one parameter all the complex factors which were difficult to determine from first principles.

This hypothesis and the values of the constants $\{C(r) \mid r \in R\}$ can be used to construct a vector stochastic process $\{X_s(t) \mid s \in S, t \geq 0\}$, where the trajectories of the stochastic process $X_s(t)$ are thought to be similar to the quantities $\mathcal{X}_s(\psi_t(\phi))$. In the stochastic model the occurrence of a chemical reaction in the future is independent of events in the past; it only depends on the availability of the reactants at the current state and the constant $C(r)$. This is clearly not true in the process $\{\mathcal{X}_s(\psi_t(\phi))\}$, although there may be only weak dependencies. Unfortunately almost no rigorous mathematical results are known in this direction that would support our underlying hypothesis. The details of the construction of this stochastic process are given in Chapter 5.

We have given the function $W(\mathcal{X}, \mathcal{X}')$ which labels the oriented edge $(\mathcal{X}, \mathcal{X}')$ in the discussion of the State graph determined by the corresponding Petri net; we also give the function $C(r)$ which specifies the reaction rate for reaction r in the discussion of Petri net. Now if we let $C(r)$ to be defined as (3.1), and define $W(\mathcal{X}, \mathcal{X}')$ as (see section 4.1):

$$W(\mathcal{X}', \mathcal{X})dt = \text{to the first order in } dt, \text{ the probability that a reaction } r \text{ would occur}$$

to reach state $\mathcal{X}$ in the time interval $(t, t + dt)$, given state $\mathcal{X}'$ at

time t,

then we can relate these two functions as follows:

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r) = C(r)V^{1-l_r} \prod_{(s,r)\in A} \binom{\mathcal{X}(s)}{I(s,r)}. \tag{3.2}$$

Next we will explain how the constant $C(r)$ is measured or estimated. Macroscopic rates of chemical reactions are expressed in terms of the concentration $\mathcal{X}_s(\phi)/V$ of molecules of the species s in the volume V. The **Reaction-Rate Constant** of reaction r, which is denoted by K_r, forms the basis for the deterministic approach to chemical kinetics. When the number $\mathcal{X}_s(\phi)$ is large compared to the typical variations

$$\mathcal{X}_s(\phi) - \frac{1}{2\,dt} \int_{-dt}^{dt} \mathcal{X}_s(\psi_t(\phi))\, dt$$

for dt small (but again not too small), then we may approximate $\mathcal{X}_s(\psi_t(\phi))/V$ by a continuous function of time: $Y_s(t)$. It is postulated that in such situations these functions satisfy a system of ordinary differential equations called the **Reaction Rate equation**:

$$Y_s'(t) = \sum_{(r,s)\in B} K_r O(r,s) \prod_{(s',r)\in A} Y_{s'}(t)^{I(s',r)} - \sum_{(s,r)\in A} K_r I(s,r) \prod_{(s',r)\in A} Y_{s'}(t)^{I(s',r)}. \tag{3.3}$$

The reaction-rate constants K_r can usually be measured for each reaction r separately, and then these rates are assumed to continue to hold in combination with other reactions. The mathematical relation between this ordinary differential equation model and the stochastic model based on the above hypothesis is elucidated in the case where all the reactions involved are reversible in ([15]): if

$$C(r) = K_r \prod_{(s,r)\in A} I(s,r)!, \tag{3.4}$$

then $\forall T > 0$, $\forall \epsilon > 0$, and $\forall s \in S$,

$$\lim_{V\to\infty} P\{\sup_{t\le T} |V^{-1}X_s(t) - Y_s(t)| > \epsilon\} = 0,$$

21

where $X_s(t)$ is the stochastic process whose jumping rates are determined by the constants $C(r)$, and $Y_s(t)$ solves the system of ordinary differential equations (3.3), and where $Y_s(0) = \lim\limits_{V \to \infty} V^{-1} X_s(0)$ with probability one.

Measurements of K_r are done on a laboratory scale whereas the stochastic model is employed for individual cells, so comparatively speaking K_r is measured in a context of nearly infinite volume. The assumption that $C(r)$ is independent of volume allows an extrapolation to a small volume situation. Since it is much easier to obtain the value of K_r than $C(r)$ of reaction r, we can use (3.4) to find the value of $C(r)$ from K_r.

The relation (3.4) shows that the Reaction Rate equation (3.3) is completely determined by the labeled Petri net.

We give a couple of simple examples which specialize the relation between $C(r)$ and K_r by (3.4) in the following.

Suppose a chemical system with volume V contains species s_1 and s_2 which undergo the reaction r:

$$s_1 \to 2s_2,$$
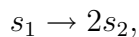
we have $l_r = 1$, and $C(r) = K_r$.

Consider a reaction r', the reverse reaction of r:

$$2s_2 \to s_1,$$
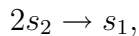
we have $l_r = 2$, and $C(r) = 2K_r$.

From above we can see, from a numerical point of view, the stochastic reaction constant and the reaction-rate constant only differ at most by a simple constant factor. But from a theoretical point of view, the difference between them is much

more complicated related to the basic conceptual differences between the stochastic and deterministic approaches ([9]).

The deterministic approach is good to adopt when the number of molecules of every species in the system is large and the fluctuation is relatively small, where we consider the system is determined by molecule concentrations that vary continuously. But for the chemical system with low molecule concentrations and slow reaction rates, a stochastic description of the temporal behavior of the reaction system seems to have a more rigorous physical basis ([9]). The deterministic approach may not apply even at higher concentrations and reaction rates when the system experiences large and rapid transitions. More discussion about the deterministic and stochastic models of chemical system can be found in ([6]) and ([15]).

On the cellular level the discrete property of the system is important due to the small number of molecules. Changes in concentration are not continuous but fluctuating. Also simulation shows that proteins are produced from an activated promoter in short bursts of variable numbers which occur at random time intervals ([16]). Hence we will adopt the stochastic approach to describe the time evolution of a genetic system involving the processes of gene expression. ([3]) and ([13]) give us a systematical discussion about how to use the deterministic and stochastic approaches to model and simulate the genetic and biochemical networks.

3.3   Markov Processes and Birth and Death Processes

To understand the whole trajectory of the state of a chemical system, we need a stochastic process.

A **Stochastic Process** is a collection of random variables $\{X(t), t \in \Gamma\}$ indexed by Γ defined on a common probability space $(\mathcal{T}, \mathcal{M}, P)$. A process is said to have

continuous time if $\Gamma = [0, \infty)$. A **Markov Process** is a stochastic process with the property that the future behavior of the process depends on the past only through its most recent value ([18]). In chemical systems, the Markov property translates into the requirement that the reactions rates depend only on the current state of the system ([20]).

To make sense of this statement we recall that a conditional probability is defined for any events $A, B \in \mathcal{M}$, where $P(A) > 0$, by the rule:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}.$$

The random variables $X(t)$ can be used to define interesting events. So now we can define a Markov process in formal terms: a process $\{X(t), t \geq 0\}$ with a countable state space $\mathcal{V}$ is said to be **Markovian** if, given $t_1 < t_2 < ... < t_n < t_{n+1}$, and $\mathcal{X}_1, ..., \mathcal{X}_{n+1} \in \mathcal{V}$,,

$$P(X(t_{n+1}) = \mathcal{X}_{n+1} \mid X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2, ..., X(t_n) = \mathcal{X}_n)$$

$$= P(X(t_{n+1}) = \mathcal{X}_{n+1} \mid X(t_n) = \mathcal{X}_n)$$

whenever $P(X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2, ..., X(t_n) = \mathcal{X}_n) > 0$.

It means we can define the time evolution of the process in terms of its conditional probabilities. For example, suppose $t_1 < t_2 < t_3$. Then from the Markov property, we have

$$P(X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2, X(t_3) = \mathcal{X}_3)$$

$$= P(X(t_2) = \mathcal{X}_2, X(t_3) = \mathcal{X}_3 \mid X(t_1) = \mathcal{X}_1) \cdot P(X(t_1) = \mathcal{X}_1)$$

$$= P(X(t_3) = \mathcal{X}_3 \mid X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2)$$

$$\cdot P(X(t_2) = \mathcal{X}_2 \mid X(t_1) = \mathcal{X}_1) \cdot P(X(t_1) = \mathcal{X}_1)$$

$$= P(X(t_3) = \mathcal{X}_3 \mid X(t_2) = \mathcal{X}_2) \cdot P(X(t_2) = \mathcal{X}_2 \mid X(t_1) = \mathcal{X}_1)$$

$$\cdot P(X(t_1) = \mathcal{X}_1)$$

whenever $P(X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2) > 0$.

Similarly, an arbitrary joint probability can be simply expressed as

$$P(X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2, ..., X(t_n) = \mathcal{X}_n)$$
$$= \left[\prod_{j=1}^{n-1} P(X(t_{j+1}) = \mathcal{X}_{j+1} \mid X(t_j) = \mathcal{X}_j) \right] \cdot P(X(t_1) = \mathcal{X}_1)$$

whenever $P(X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2, ..., X(t_{n-1}) = \mathcal{X}_{n-1}) > 0$.

For $A, B \in \mathcal{M}$, define

$$P^*(B \mid A) = \begin{cases} \frac{P(B \cap A)}{P(A)} & \text{when } P(A) > 0, \\ \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that the state space $\mathcal{V}$ of the Markov process is countable. Using the Markovian property of the process, it can be derived that for $t_1 < t_2 < t_3$, when $P(X(t_1) = \mathcal{X}_1) > 0$,

$$P(X(t_3) = \mathcal{X}_3 \mid X(t_1) = \mathcal{X}_1)$$
$$= \sum_{\mathcal{X}_2 \in \mathcal{V}} P(X(t_2) = \mathcal{X}_2, X(t_3) = \mathcal{X}_3 \mid X(t_1) = \mathcal{X}_1)$$
$$= \sum_{\mathcal{X}_2 \in \mathcal{V}} P^*(X(t_3) = \mathcal{X}_3 \mid X(t_1) = \mathcal{X}_1, X(t_2) = \mathcal{X}_2) P(X(t_2) = \mathcal{X}_2 \mid X(t_1) = \mathcal{X}_1)$$
$$= \sum_{\mathcal{X}_2 \in \mathcal{V}} P^*(X(t_3) = \mathcal{X}_3 \mid X(t_2) = \mathcal{X}_2) \cdot P(X(t_2) = \mathcal{X}_2 \mid X(t_1) = \mathcal{X}_1).$$

The equation

$$P(X(t_3) = \mathcal{X}_3 \mid X(t_1) = \mathcal{X}_1)$$
$$= \sum_{\mathcal{X}_2 \in \mathcal{V}} P^*(X(t_3) = \mathcal{X}_3 \mid X(t_2) = \mathcal{X}_2) \cdot P(X(t_2) = \mathcal{X}_2 \mid X(t_1) = \mathcal{X}_1)$$

$$(3.5)$$

holding whenever $P(X(t_1) = \mathcal{X}_1) > 0$ is called the **Chapman-Kolmogorov equation** in the countable state space.

A **Birth and Death Process** is a particular type of continuous time, discrete state Markov process. We deal with a family of random variables $\{X(t), t \geq 0\}$ whose possible values are nonnegative integers, i.e. $\mathcal{V} = \mathbb{N} \cup \{0\}$. If at time t the process is in state n it may, after a random waiting time, move to either of the neighboring states $n + 1$ because of the occurrence of the birth reaction or $n - 1$ because of the occurrence of the death reaction.

The associated State graph of the Birth and Death process is given in Fig 3.2, where $i \geq 1$ represents a state of the process, and there is no largest state, i.e. $k \to \infty$.
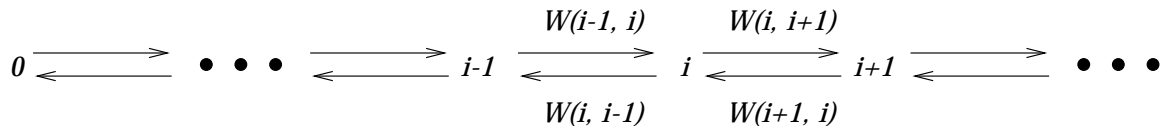
Figure 3.2: State Graph for the Birth and Death Process with $k \to \infty$

The transition probability of a Birth and Death process is denoted by

$$P_{ij}(t) = P(X(t + s) = j \mid X(s) = i), \qquad i, j \in \mathcal{V} = \{0, 1, 2, ...\}$$

when $P(X(s) = i) > 0$. Here we assume that $\{X(t)\}$ has the stationary transition property, i.e. $P(X(t + s) = j \mid X(s) = i)$ is independent of $s \geq 0$.

When $t = 0$, we have

$$P_{ij}(0) = 0, \qquad i \neq j$$
$$P_{ij}(0) = 1, \qquad i = j \tag{3.6}$$

In addition we assume $P_{ij}(t)$ satisfies the following properties when $P_{ij}(t)$ is well-defined:

(1) $\mu_0 = 0, \lambda_0 > 0$, and $\mu_i > 0, \lambda_i > 0$ for $i \geq 1$.

(2) $P_{i,i+1}(h) = \lambda_i h + o(h)$ as $h \downarrow 0$, $\forall i \geq 0$.

(3) $P_{i,i-1}(h) = \mu_i h + o(h)$ as $h \downarrow 0$, $\forall i \geq 1$.

(4) $P_{ii}(h) = 1 - (\lambda_i + \mu_i)h + o(h)$ as $h \downarrow 0$, $\forall i \geq 0$.

(5) $\forall j \geq 0$, $\exists \epsilon > 0$, $\sup\{\dfrac{P_{ij}(h)}{h} \mid 0 < h < \epsilon, i \geq 0, i \neq j-1, j, j+1\} < \infty$.

The parameters λ_i and μ_i are called the infinitesimal birth and death rates respectively. Since the $P_{ij}(t)$ are probabilities, we have $P_{ij}(t) \geq 0$ and $\sum_{j=0}^{\infty} P_{ij}(t) = 1$.

By postulates 2 and 3, if the process starts from state i, then in a small time interval the probabilities of the population increasing or decreasing by 1 are essentially proportional to the length of the interval. Given a transition from state i occurs at time t, the probability that this transition is to state $i+1$ is $\lambda_i(\lambda_i + \mu_i)^{-1}$, and to state $i-1$ is $\mu_i(\lambda_i + \mu_i)^{-1}$. The postulates and properties about the Birth and Death processes are adapted from ([14]).

The following equations

$$\frac{d}{dt}P_{ij}(t) = \lambda_{j-1}P_{i,j-1}(t) + \mu_{j+1}P_{i,j+1}(t) - (\lambda_j + \mu_j)P_{ij}(t) \quad \text{for } j \geq 1, \qquad (3.7)$$

and

$$\frac{d}{dt}P_{i0}(t) = -\lambda_0 P_{i0}(t) + \mu_1 P_{i1}(t), \qquad (3.8)$$

when $P_{ij}(t)$ is well defined are called the **forward Kolmogorov differential equation** for Birth and Death processes with initial conditions (3.6).

In the following, we will compare a Markov process which describes the time behavior of a chemical system with a Birth and Death process.

Let $\mathcal{X}$ be a state of the system, and $\mathcal{X}(s)$ be the number of molecules of species s, where $\mathcal{X} : S \longrightarrow \mathbb{N} \cup \{0\}$, then $\mathcal{X} = (\mathcal{X}(s))_{s \in S}$. $\mathcal{X}(s)$ decreases as a reaction occurs

in which s is involved as a reactant, and increases when a reaction occurs in which s is a product. It is easy to see that for each chemical species, the fluctuation of its molecule number is similar to a birth and death process in the sense that it can either increase or decrease. We can also observe two differences between the process of a species involving in reactions of the chemical system we are concerning with and the Birth and Death process. First, suppose the current state of a species s is $\mathcal{X}(s)$, when the next reaction involving this species occurs at a random later time, it may jump directly to the states $\mathcal{X}(s) + j$ or $\mathcal{X}(s) - j$, where $j > 1$. While in the birth and death process, the next state of the species could only be $\mathcal{X}(s) + 1$ or $\mathcal{X}(s) - 1$. Second, the transition probabilities of s depends not only on its own present state but also the present states of other species involved in the same reactions which s is involved in. But in the Birth and Death process, the transition probabilities of one species only depend on its own present state.

Let us consider the following chemical system involving the following two reactions r_1 and r_2 as an example:

$$r_1 : A \longrightarrow 2B,$$

$$r_2 : B + C \longrightarrow D.$$
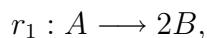
The Petri net representation for this system is shown in Fig 3.3.
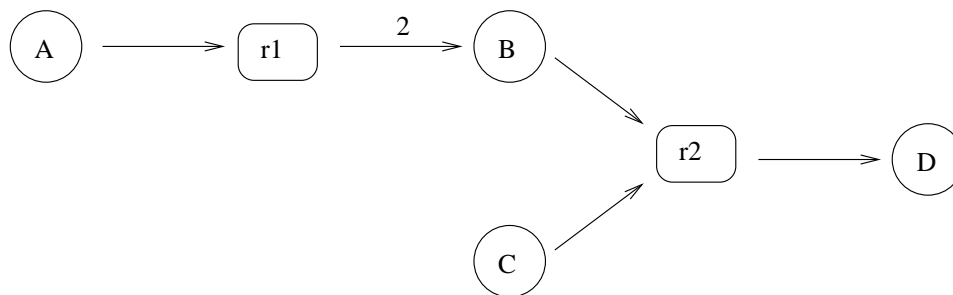
Figure 3.3: Petri Net for the System involving Reactions r_1 and r_2.

The changes in the population structure of B are effected as follows: if at time t the number of B molecules is i it may, after a random waiting time, move to another state $i+2$, as 2 molecules of the product B are 'born' when the first reaction occurs. At a random later time, the state of B may decrease from $i+2$ to $i+1$, as 1 molecule of B 'dies' when the second reaction occurs.

Therefore, to describe the time evolution of a chemical system, we must deal with the state of the whole system instead of each species separately.

Chapter 4

Master Equation Approach for Chemical Systems

4.1 Derivation of the Master Equation from a Markov Process

In this section we are concerned with an abstract directed graph $G = (\mathcal{V}, E)$ with a countable number of vertices and no loops. The main example we have in mind are the State graphs associated to Petri nets. Suppose a function $W : \mathcal{V} \times \mathcal{V} \to (0, \infty)$ is given, then for two vertices v' and v, where $(v', v) \in E$, we may consider the oriented edge (v', v) to be labeled by the number $W(v', v)$.

The Master equation associated with the above labeled directed graph has the following form:

$$\frac{d}{dt} f_v(t) = \sum_{(v',v) \in E} W(v', v) f_{v'}(t) - [\sum_{(v,v'') \in E} W(v, v'')] f_v(t), \qquad (4.1)$$

where for each $v \in \mathcal{V}$, $f_v : \mathbb{R} \to \mathbb{R}$ is a $\mathbf{C}^1$ function.

The following is an example of the Master equation for the Birth and Death process. The system contains only one species s and two simple chemical reactions:

$$r_1(Birth) : \qquad \emptyset \to s, \qquad (4.2)$$

$$r_2(Death) : \qquad s \to \emptyset, \qquad (4.3)$$

i.e., for $j \geq 1$,

$$r_1: \quad j-1 \xrightarrow{W(j-1,j)} j,$$

$$r_2: \quad j+1 \xrightarrow{W(j+1,j)} j,$$

where $W(j, j+1) = \lambda_j$ and $W(j, j-1) = \mu_j$.

Let $f_j(t)$ denote the probability that the state of the system is state j at time t, then the **Birth-Death Master Equation** ([7]) takes the form:

$$\frac{d}{dt} f_j(t) = W(j-1, j) f_{j-1}(t) + W(j+1, j) f_{j+1}(t) - [W(j, j+1) + W(j, j-1)] f_j(t).$$
$$(4.4)$$

When $j = 0$,

$$\frac{d}{dt} f_0(t) = W(1, 0) f_1(t) - W(0, 1) f_0(t). \tag{4.5}$$

Another example is the Master equation for a chemical system with volume V which contains chemical species $s \in S$ interreacting through chemical reactions $r \in R$.

We know that given a state $\mathcal{X}' \in \mathcal{V}$, a reaction r determines a possible transition from $\mathcal{X}'$ to another state $\mathcal{X} \in \mathcal{V}$. Define $W(\mathcal{X}', \mathcal{X})$ to be the probability of direct transition from state $\mathcal{X}'$ to $\mathcal{X}$ through some reaction r per unit time ([20]), then

$$\begin{aligned} W(\mathcal{X}', \mathcal{X}) dt = {}&\text{to the first order in } dt, \text{ the probability that a reaction } r \text{ would occur} \\ &\text{to reach state } \mathcal{X} \text{ in the time interval } (t, t+dt), \text{ given state } \mathcal{X}' \text{ at} \\ &\text{time } t. \end{aligned} \tag{4.6}$$

Observe that given the state $\mathcal{X}'$, if $(\mathcal{X}', \mathcal{X}) \notin E$, then $W(\mathcal{X}', \mathcal{X}) = 0$.

Hence we can represent a labeled edge $(\mathcal{X}', \mathcal{X})$ of the State graph $G = (\mathcal{V}, E)$ arising from the reaction r as follows:

$$\mathcal{X}' \xrightarrow{W(\mathcal{X}', \mathcal{X})} \mathcal{X} \tag{4.7}$$

where $\mathcal{X} = \mathcal{X}' - \mathbf{n}_r$.

Let $f_{\mathcal{X}}(t)$ denote the probability that the state of the system is $\mathcal{X}$ at time t, i.e. $f_{\mathcal{X}}(t) = P(X(t) = \mathcal{X})$. Then the corresponding Master equation for the chemical system is:

$$\frac{d}{dt} f_{\mathcal{X}}(t) = \sum_{(\mathcal{X}',\mathcal{X}) \in E} W(\mathcal{X}', \mathcal{X}) f_{\mathcal{X}'}(t) - [\sum_{(\mathcal{X},\mathcal{X}'') \in E} W(\mathcal{X}, \mathcal{X}'')] f_{\mathcal{X}}(t).$$

From above, we see that the Master equation can be used to compute the time dependent probability distribution of the state of a chemical system.

Next, we will derive the Master equation associated with the abstract directed graph $G = (\mathcal{V}, E)$. Assume now there exist a Markov process $\{X(t), t \geq 0\}$ on the probability space $(\mathcal{T}, \mathcal{M}, P)$ which has values in the set $\mathcal{V}$, which we will continue to call the set of states. Such a Markov process is a continuous time process with stationary transition probabilities. Thus $X(t) : [0, \infty) \longrightarrow \mathcal{V} \ \forall t \geq 0$. The traditional stochastic approach to compute the probabilities $f_v(t)$ of a particular state v at a given time t is to set up and solve the corresponding Master Equation for the process, which is built on the Markov model foundation. This derivation of the Master equation is adapted from ([9]).

Assume the Markov process satisfies the following postulates:
for $P(X(t) = v') > 0$,

(1) If $(v', v) \in E$, then $P(X(t+dt) = v \mid X(t) = v') = W(v', v)dt + o(dt)$ as $dt \downarrow 0$;

(2) If $v \in \mathcal{V}$, then $P(X(t+dt) = v \mid X(t) = v) = 1 - \sum_{v'' \in \mathcal{V}, (v,v'') \in E} W(v, v'')dt + o(dt)$ as $dt \downarrow 0$;

(3) $\forall v \in \mathcal{V}, \exists \epsilon > 0, \sup\{\frac{P(X(t+dt) = v \mid X(t) = v')}{dt} \mid 0 < dt < \epsilon, v' \in \mathcal{V}, (v', v) \notin E \cup \Delta\} < \infty$.

Here E is the set of oriented edges in G, and $\Delta = \{(v,v) \mid v \in \mathcal{V}\} \subset \mathcal{V} \times \mathcal{V}$.

Let $P(X(t) = v \mid X(0) = v_0)$ be the probability that the state of the process is v at time t given the initial state is v_0 at time 0, then

$$
\begin{aligned}
f_{\mathcal{X}}(t) &= \sum_{v_0 \in \mathcal{V}} P^*(X(t) = v \mid X(0) = v_0) \cdot f_{v_0}(0) \\
&= \sum_{v_0 \in \mathcal{V}} P^*(X(t) = v \mid X(0) = v_0) \cdot P(X(0) = v_0).
\end{aligned}
\tag{4.8}
$$

In the following we will show the conditional probability $P(X(t) = v \mid X(0) = v_0)$ satisfies the Master equation based on these 3 postulates when $P(X(0) = v_0) > 0$.

To reach state v during time interval $(t, t+dt)$ from the initial state v_0 at time $t_0 = 0$, there are several possibilities: the process arrives at state v through a transition during $(t, t + dt)$ from an appropriate once-removed state v' at time t; the process reaches state v at time t, and remains in that state in $(t, t + dt)$; the process reaches state v through two or more transitions during $(t, t+dt)$. From Chapman-Kolmogorov equation (3.5), we have if $P(X(0) = v_0) > 0$, then

$$
\begin{aligned}
&P(X(t + dt) = v \mid X(0) = v_0) \\
&= \sum_{v' \in \mathcal{V}} P^*(X(t + dt) = v \mid X(t) = v') \cdot P(X(t) = v' \mid X(0) = v_0) \\
&= \sum_{v' \in \mathcal{V}, (v',v) \in E} P^*(X(t + dt) = v \mid X(t) = v') \cdot P(X(t) = v' \mid X(0) = v_0) \\
&\quad + P^*(X(t + dt) = v \mid X(t) = v) \cdot P(X(t) = v \mid X(0) = v_0) \\
&\quad + \sum_{v' \in \mathcal{V}, (v',v) \notin E \cup \Delta} P^*(X(t + dt) = v \mid X(t) = v') \cdot P(X(t) = v' \mid X(0) = v_0).
\end{aligned}
$$

When $P(X(t) = v') > 0$, by postulates 1, 2, and

$$\sum_{v \in \mathcal{V}} P(X(t + dt) = v \mid X(t) = v') = \sum_{v \in \mathcal{V}, (v',v) \in E} P(X(t + dt) = v \mid X(t) = v')$$
$$+ P(X(t + dt) = v' \mid X(t) = v')$$
$$+ \sum_{v \in \mathcal{V}, (v',v) \notin E \cup \Delta} P(X(t + dt) = v \mid X(t) = v')$$
$$= 1,$$

we can easily obtain that, for any $v' \in \mathcal{V}$ with $P(X(t) = v') > 0$,

$$0 \leq \sum_{v \in \mathcal{V}, (v',v) \notin E \cup \Delta} P(X(t + dt) = v \mid X(t) = v') = o(dt).$$

Hence $\forall v' \in \mathcal{V}$ with $P(X(t) = v') > 0$, $\forall v \in \mathcal{V}$, s.t. $(v', v) \notin E \cup \Delta$,

$$P(X(t + dt) = v \mid X(t) = v') = o(dt), \tag{4.9}$$

i.e.

$$\lim_{dt \to 0} \frac{P(X(t + dt) = v \mid X(t) = v')}{dt} = 0.$$

By postulate 3 we have, for any $v \in \mathcal{V}$, $\exists C < \infty$, $\exists \epsilon > 0$, s.t. $\forall\ 0 < dt < \epsilon$, $\forall v' \in \mathcal{V}$ with $P(X(t) = v') > 0$, and $(v', v) \notin E \cup \Delta$,

$$\frac{P(X(t + dt) = v \mid X(t) = v')}{dt} < C,$$

thus

$$\sum_{v', (v',v) \notin E \cup \Delta} \frac{P(X(t + dt) = v \mid X(t) = v')}{dt} \cdot P(X(t) = v' \mid X(0) = v_0)$$
$$\leq C \cdot \sum_{v', (v',v) \notin E \cup \Delta} P(X(t) = v' \mid X(0) = v_0)$$
$$\leq C.$$

When $P(X(t) = v') = 0$, we have $P^*(X(t + dt) = v \mid X(t) = v') = 0$ by the definition of P^*, i.e. $\forall v' \in \mathcal{V}$ with $P(X(t) = v') = 0$,

$$P^*(X(t + dt) = v \mid X(t) = v') \cdot P(X(t) = v' \mid X(0) = v_0) = 0. \tag{4.10}$$

34

Hence by the Dominated Convergence theorem, we obtain

$$
\lim_{dt \to 0} \sum_{v' \in \mathcal{V}, (v',v) \notin E \cup \Delta} \frac{P^*(X(t+dt) = v \mid X(t) = v')}{dt} \cdot P(X(t) = v' \mid X(0) = v_0)
$$

$$
= \sum_{v' \in \mathcal{V}, (v',v) \notin E \cup \Delta} \lim_{dt \to 0} \frac{P^*(X(t+dt) = v \mid X(t) = v')}{dt} \cdot P(X(t) = v' \mid X(0) = v_0)
$$

$$
= 0,
$$

i.e.,

$$
\sum_{v' \in \mathcal{V}, (v',v) \notin E \cup \Delta} P^*(X(t+dt) = v \mid X(t) = v') \cdot P(X(t) = v' \mid X(0) = v_0) = o(dt).
$$

(4.11)

By postulates 1, 2, (4.10) and (4.11), we have

$$
P(X(t+dt) = v \mid X(0) = v_0)
$$

$$
= \sum_{v' \in \mathcal{V}, (v',v) \in E} [W(v', v)dt + o(dt)] \cdot P(X(t) = v' \mid X(0) = v_0)
$$

$$
+ [1 - \sum_{v' \in \mathcal{V}, (v,v'') \in E} W(v, v'')dt + o(dt)] \cdot P(X(t) = v \mid X(0) = v_0)
$$

$$
+ o(dt).
$$

(4.12)

We obtain directly from (4.12):

$$
\frac{d}{dt} P(X(t) = v \mid X(t_0) = v_0)
$$

$$
= \sum_{v' \in \mathcal{V}, (v',v) \in E} W(v', v) \cdot P(X(t) = v' \mid X(0) = v_0)
$$

$$
- [\sum_{v'' \in \mathcal{V}, (v,v'') \in E} W(v, v'')] \cdot P(X(t) = v \mid X(0) = v_0).
$$

(4.13)

We have derived that the conditional probability $P(X(t) = v \mid X(0) = v_0)$ satisfies the Master equation under the assumption that $P(X(0) = v_0) > 0$. For $P^*(X(t) = v \mid X(0) = v_0)$ when $P(X(0) = v_0) = 0$, again it has been assigned value 0 by the definition of P^*. Then we will show that $f_v(t)$ also satisfies the Master equation. By (4.8) and (4.13) we have $\forall v \in \mathcal{V}$,

$$\frac{d}{dt}f_v(t) = \frac{d}{dt}\sum_{v_0 \in \mathcal{V}} P^*(X(t) = v \mid X(0) = v_0)f_{v_0}(0)$$

$$= \sum_{v_0 \in \mathcal{V}}[\frac{d}{dt}P^*(X(t) = v \mid X(0) = v_0)]f_{v_0}(0)$$

$$= \sum_{v_0 \in \mathcal{V}}[\sum_{(v',v) \in E} W(v',v)P^*(X(t) = v' \mid X(0) = v_0)]f_{v_0}(0)$$

$$- \sum_{v_0 \in \mathcal{V}}[\sum_{(v,v'') \in E} W(v,v'')]P^*(X(t) = v \mid X(0) = v_0)f_{v_0}(0)$$

$$= \sum_{(v',v) \in E} W(v',v)[\sum_{v_0 \in \mathcal{V}} P^*(X(t) = v' \mid X(0) = v_0)f_{v_0}(0)]$$

$$- \sum_{(v,v'') \in E} W(v,v'')[\sum_{v_0 \in \mathcal{V}} P^*(X(t) = v \mid X(0) = v_0)f_{v_0}(0)]$$

$$= \sum_{(v',v) \in E} W(v',v)f_{v'}(t) - [\sum_{(v,v'') \in E} W(v,v'')]f_v(t).$$

which is exactly the **Master Equation**.

4.2 Theory of the Master Equation

In this section we will derive the major properties of solutions of the Master Equation associated with the abstract labeled directed graph $G = (\mathcal{V}, E)$ with a finite number of vertices and no loops.

For $v \in \mathcal{V}$, define $\alpha(v) = \sum_{(v,v'') \in E} W(v,v'')$ and $\alpha = \max_{v \in \mathcal{V}} \alpha(v)$, then $\alpha(v) \geq 0$ and $\alpha > 0$. And

$$\frac{d}{dt}f_v(t) = \sum_{v' \in V} A_{vv'}f_{v'}(t),$$

where A is a matrix whose entries are

$$A_{vv'} = \begin{cases} -\alpha(v) & \text{if } v' = v, \\ W(v',v) & \text{if } (v',v) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that the matrix A has the following properties: each row and each

column has $|R| + 1$ nonzero entries at most; each column of A sums to zero since

$$\sum_{v \in \mathcal{V}} A_{vv'} = \sum_{(v',v) \in E} W(v', v) - \alpha(v')$$

$$= \sum_{(v',v) \in E} W(v', v) - \sum_{(v',v) \in E} W(v', v)$$

$$= 0.$$

Observe that the number of entries in each column may not be equal because of the boundary conditions, but the statement above still holds true by the way we defined $\alpha(v)$.

Since $\mathcal{V}$ is finite, we have $\max_{(v',v) \in E} W(v', v) < \infty$. Thus the solution of the Master Equation (4.1) exists, where $f(t) = e^{At} f(0)$, where $f(t)$ is the vector $(f_v(t))_{v \in \mathcal{V}}$.

Theorem 4.1: $\sum_{v \in \mathcal{V}} f_v(t) \equiv C \ \forall t \in \mathbb{R}$. In particular, if $\sum_{v \in \mathcal{V}} f_v(0) = 1$, then $\sum_{v \in \mathcal{V}} f_v(t) = 1 \ \forall t \in \mathbb{R}$.

Proof: We have

$$\frac{d}{dt} \sum_{v \in \mathcal{V}} f_v(t) = \sum_{v \in \mathcal{V}} [\sum_{(v',v) \in E} W(v', v) f_{v'}(t)] - \sum_{v \in \mathcal{V}} [\sum_{(v,v'') \in E} W(v, v'')] f_v(t)$$

$$= \sum_{e \in E} W(e) f_{\pi_1(e)}(t) - \sum_{e \in E} W(e) f_{\pi_1(e)}(t)$$

$$= 0,$$

where $\pi_1(v, v') = v$. This implies that $\sum_{v \in \mathcal{V}} f_v(t)$ is independent of t. $\square$

Theorem 4.2: The solution f_v of Master equation satisfies the non-negative property, i.e., if $\forall \ v \in \mathcal{V}$, $f_v(0) \geq 0$, then $f_v(t) \geq 0$, $\forall \ t > 0 \ \forall \ v \in \mathcal{V}$.

Proof: Since

$$\frac{d}{dt} f_v(t) + \alpha(v) f_v(t) = \sum_{(v',v) \in E} W(v', v) f_{v'}(t),$$

i.e.,

$$\frac{d}{dt}[e^{\alpha(v)t}f_v(t)] = \sum_{(v',v)\in E} W(v',v)f_{v'}(t)e^{\alpha(v)t},$$

then

$$e^{\alpha(v)t}f_v(t) - f_v(0) = \sum_{(v',v)\in E} W(v',v)\int_0^t f_{v'}(s)e^{\alpha(v)s}ds,$$

we obtain

$$f_v(t) = f_v(0)e^{-\alpha(v)t} + \sum_{(v',v)\in E} W(v',v)\int_0^t f_{v'}(s)e^{-\alpha(v)(t-s)}ds. \tag{4.14}$$

Define $g(t) = \min\{0, \min_{v\in\mathcal{V}} f_v(t)\}$, $\beta(v) = \sum_{(v',v)\in E} W(v',v)$, and $\beta = \max_{v\in\mathcal{V}} \beta(v)$, then $g(t) \le 0$ and $g(t) \le f_v(t)$ for all t, $\beta(v) \ge 0$, and $\beta > 0$. We have

$$f_v(t) \ge g(0)e^{-\alpha(v)t} + \sum_{(v',v)\in E} W(v',v)\int_0^t g(s)e^{-\alpha(v)(t-s)}ds$$

$$\ge g(0) + \beta(v)\int_0^t g(s)ds$$

$$\ge g(0) + \beta\int_0^t g(s)ds$$

since $e^{-\alpha(v)t} \le 1$. Hence

$$\min_{v\in\mathcal{V}} f_v(t) \ge g(0) + \beta\int_0^t g(s)ds,$$

i.e.,

$$g(t) \ge g(0) + \beta\int_0^t g(s)ds. \tag{4.15}$$

Let $G(t) = \int_0^t g(s)ds$. $g(t)$ is continuous, so G is $\mathbf{C}^1$. Therefore,

$$G'(t) - \beta G(t) \ge g(0),$$

i.e.,

$$\frac{d}{dt}[e^{-\beta t}G(t)] \ge g(0)e^{-\beta t}.$$

Hence

$$e^{-\beta t}G(t) - G(0) \geq g(0)\int_0^t e^{-\beta t}ds,$$

$$e^{-\beta t}G(t) \geq \frac{g(0)}{\beta}(1 - e^{-\beta t}),$$

$$G(t) \geq \frac{g(0)}{\beta}(e^{\beta t} - 1).$$

We obtain from (4.15)

$$\frac{g(t) - g(0)}{\beta} \geq \frac{g(0)}{\beta}(e^{\beta t} - 1),$$

i.e.,

$$g(t) \geq g(0)e^{\beta t}. \tag{4.16}$$

The meaning of (4.16) is that if $g(0) = 0$, then $g(t) \equiv 0 \ \forall t \geq 0$. That is, if $\forall v \in \mathcal{V}$, $f_v(0) \geq 0$, then $f_v(t) \geq 0 \ \forall \ t \geq 0 \ \forall v \in \mathcal{V}$. $\square$

For each $v \in \mathcal{V}$, define

$I_v = \{v' \in \mathcal{V} \mid$ there is an oriented path of length at least 1 in $(\mathcal{V}, E)$ from v' to $v\}$,

and given $f_v(0) \geq 0 \ \forall v \in \mathcal{V}$, define

$$V_0 = \{v \in \mathcal{V} \mid \exists v' \in I_v, f_{v'}(0) > 0\} \cup \{v \in \mathcal{V} \mid f_v(0) > 0\}.$$

Thus, $v \in \mathcal{V} \setminus V_0$ if and only if $f_v(0) = 0$ and $\forall v' \in I_v, f_{v'}(0) = 0$.

Theorem 4.3: Assume $f_v(0) \geq 0 \ \forall v \in \mathcal{V}$. If $v \in \mathcal{V} \setminus V_0$, then $f_v(t) \equiv 0 \ \forall \ t \geq 0$.

Proof: If $v \in \mathcal{V} \setminus V_0$ and $v' \in I_v$, then $f_{v'}(0) = 0$; also $\forall \ v'' \in I_{v'}, f_{v''}(0) = 0$ since $I_{v'} \subset I_v$. This implies that $v' \in \mathcal{V} \setminus V_0$. Hence by (4.14) $\forall \ v \in \mathcal{V} \setminus V_0$,

$$f_v(t) = \sum_{(v',v)\in E} W(v',v)\int_0^t f_{v'}(s)e^{-(t-s)\alpha(v)}ds$$

$$\leq \sum_{(v',v)\in E} W(v',v)\int_0^t [\max_{v''\in\mathcal{V}\setminus V_0} f_{v''}(s)]e^{-(t-s)\alpha(v)}ds$$

$$\leq \beta \int_0^t h(s)ds,$$

where $h(s) = \max\limits_{v'' \in \mathcal{V} \setminus V_0} f_{v''}(s) \geq 0$. Then

$$h(t) \leq \beta \int_0^t h(s)ds. \tag{4.17}$$

Let $H(t) = \int_0^t h(s)ds$, then $H'(t) = h(t)$ since $h(t)$ is continuous, and

$$H'(t) - \beta H(t) \leq 0,$$
$$\frac{d}{dt}[H(t)e^{-\beta t}] \leq 0,$$
$$H(t)e^{-\beta t} - H(0) \leq 0,$$
$$H(t) \leq 0.$$

By (4.17), we have $h(t) \leq 0$.

Hence, $h(t) = 0$, i.e., $f_v(t) \equiv 0 \ \forall \ t \geq 0, \ \forall \ v \in \mathcal{V} \setminus V_0$. $\square$

Theorem 4.4: Assume $f_v(0) \geq 0 \ \forall v \in \mathcal{V}$. If $v \in V_0$, then $f_v(t) > 0 \ \forall \ t > 0$.

Proof: Define $V_1 = \{v \in V_0 \mid \exists t > 0, f_v(t) = 0\}$. Then it suffices to show that V_1 is an empty set. We prove it by contradiction.

$\forall v \in V_1$, we have $\exists t > 0$ s.t. $f_v(t) = 0$. Hence by (4.14), we have

$$f_v(0) = 0 \tag{4.18}$$

and

$$f_{v'}(s) = 0 \quad \forall (v', v) \in E \ \forall 0 \leq s \leq t \tag{4.19}$$

$\forall v \in V_1$, Since $v \in V_0$, (4.18) implies that $\exists v_1 \in I_v$, s.t. $f_{v_1}(0) > 0$. Choose $v_1 \in I_v$ s.t. $f_{v_1}(0) > 0$ and the path from v_1 to v is as short as possible. Hence from (4.19), the path from v_1 to v has length at least 2, we denote this length as $l(v)$. Then for any $v \in V_1$, we have (4.18), (4.19) and $l(v) \geq 2$ defined, where $l : V_1 \longrightarrow \{2, 3, ...\}$.

Choose $v \in V_1$ s.t. $l(v)$ is as small as possible. $l(v) \geq 2$ implies that $v_1 \in I_{v'}$ for some v' where $(v', v) \in E$. Hence this $v' \in V_0$ since $f_{v_1}(0) > 0$. Since $f_{v'}(t) = 0$ by (4.19), we have that $v' \in V_1$ and $l(v') < l(v)$. Hence we get a contradiction. $\square$

Next, we discuss the long time behavior of the nonnegative solutions of the Master equation.

We know that $A + \alpha I$ is a matrix with non-negative entries, hence by the **Perron-Frobenius theorem** ([21]), we have that there exists a nonnegative eigenvalue λ of $A + \alpha I$ such that no eigenvalue of $A + \alpha I$ has absolute value greater than λ, and corresponding to this eigenvalue λ there is at least one nonnegative eigenvector $h = \{h_v\}$, s.t,

$$\sum_{v' \in \mathcal{V}} (A_{vv'} + \alpha \delta_{vv'}) h_{v'} = \lambda h_v,$$

i.e.

$$\sum_{v' \in \mathcal{V}} A_{vv'} h_{v'} = (\lambda - \alpha) h_v.$$

Hence

$$(\lambda - \alpha) \sum_{v \in \mathcal{V}} h_v = \sum_{v \in \mathcal{V}} \sum_{v' \in \mathcal{V}} A_{vv'} h_{v'} = \sum_{v' \in \mathcal{V}} (\sum_{v \in \mathcal{V}} A_{vv'}) h_{v'} = 0.$$

Since $h_v \geq 0 \ \forall v \in \mathcal{V}$ and h is an eigenvector, we have $\sum_{v \in \mathcal{V}} h_v > 0$, thus $\lambda = \alpha$, i.e., A has a zero eigenvalue with a nonnegative eigenvector h.

Suppose λ' is any other eigenvalue of A distinct from 0, then $\lambda' + \alpha$ is an eigenvalue of $A + \alpha I$, and $\mid \lambda' + \alpha \mid \leq \alpha$, i.e.

$$\sqrt{[\mathrm{Re}(\lambda' + \alpha)]^2 + [\mathrm{Im}(\lambda' + \alpha)]^2} \leq \alpha,$$

$$[\mathrm{Re}(\lambda')]^2 + 2\alpha \mathrm{Re}(\lambda') + \alpha^2 + [\mathrm{Im}(\lambda')]^2 \leq \alpha^2,$$

since $\mathrm{Re}(\alpha) = \alpha$ and $\mathrm{Im}(\alpha) = 0$. Hence

$$\mathrm{Re}(\lambda')[\mathrm{Re}(\lambda') + 2\alpha] \leq -[\mathrm{Im}(\lambda')]^2 \leq 0.$$

41

Since $\alpha > 0$, we must have $-2\alpha \leq \text{Re}(\lambda') \leq 0$. If $\text{Re}(\lambda') = 0$, then $[\text{Im}(\lambda')]^2 = 0$; thus either $\lambda' = 0$ or $\text{Re}(\lambda') < 0$. Therefore all the eigenvalues of A distinct from 0 have negative real parts.

Suppose $f_v(0) \geq 0 \ \forall v \in \mathcal{V}$ and

$$f_v(0) = \sum_{\lambda} \sum_{J(\lambda)} \sum_{h} c_h h_v,$$

where the first summation is over all the eigenvalues λ of the matrix A, the second is over all the Jordan blocks $J(\lambda)$ belonging to the particular eigenvalue λ, the last is over the eigenvector and all generalized eigenvectors h belonging to the particular Jordan block $J(\lambda)$, and $c_h \in \mathbb{C}$ depends on the generalized eigenvector h. We say h is a $j-$eigenvector belonging to the eigenvalue λ if

$$(A - \lambda I)^{j+1} h = \mathbf{0},$$

but

$$(A - \lambda I)^j h \neq \mathbf{0}.$$

Note that $(A - \lambda I)^0 = I$.

Consider a Jordan block of size $(p+1) \times (p+1)$, then correspondingly we have one eigenvector h^0 and p generalized eigenvectors $h^1, ..., h^p$, where h^j is a $j-$eigenvector and $h^{j-1} = (A - \lambda I) h^j$. When $p = 0$, we have one eigenvector and no generalized eigenvectors. Observe that $(Ah^0, Ah^1..., Ah^p) = (\lambda h^0, h^0 + \lambda h^1, ..., h^{p-1} + \lambda h^p)$, i.e., for $0 \leq q \leq p$,

$$(A - \lambda I)^{p-q} h^p = h^q \neq \mathbf{0}.$$

Hence

$$e^{At}h^q = e^{(\lambda I + A - \lambda I)t}h^q$$

$$= e^{\lambda t}e^{(A-\lambda I)t}h^q$$

$$= e^{\lambda t}\sum_{n=0}^{q}\frac{t^n}{n!}(A-\lambda I)^n h^q$$

$$= e^{\lambda t}[h^q + t(A-\lambda I)h^q + ... + \frac{t^q}{q!}(A-\lambda I)^q h^q]$$

$$= e^{\lambda t}[h^q + th^{q-1} + ... + \frac{t^q}{q!}h^0] \tag{4.20}$$

For a general Jordan block $J(\lambda)$, let its size be $[p(J(\lambda)) + 1] \times [p(J(\lambda)) + 1]$, thus we obtain

$$f_v(t) = e^{At}f_v(0)$$

$$= \sum_{J(0)}\sum_{q=0}^{p(J(0))}c_{h^q}[(e^{At})h^q]_v + \sum_{\lambda \neq 0}\sum_{J(\lambda)}\sum_{q=0}^{p(J(\lambda))}c_{h^q}[(e^{At})h^q]_v \tag{4.21}$$

It is understood that $h^0, h^1, ..., h^{p(J(\lambda))}$ depend on the Jordan block $J(\lambda)$.

Let $\epsilon = \frac{1}{2}\min_\lambda |\text{Re}(\lambda)|$, where λ ranges over the eigenvalues of A with $\text{Re}(\lambda) < 0$. Thus

$$|e^{\lambda t}t^q| = |e^{(\text{Re}\lambda + i\text{Im}\lambda)t}||t^q|$$

$$= e^{(\text{Re}\lambda)t}|t^q|$$

$$\leq e^{-2\epsilon t}|t^q|$$

$$= e^{-\epsilon t}(e^{-\epsilon t}|t^q|)$$

$$= O(e^{-\epsilon t})$$

since $e^{-\epsilon t}|t^q| \to 0$ as $t \to \infty$. Hence the second term in (4.21) satisfies

$$\sum_{\lambda \neq 0}\sum_{J(\lambda)}\sum_{q=0}^{p(J(\lambda))}c_{h^q}[(e^{At})h^q]_v = O(e^{-\epsilon t}).$$

which tends to 0 as $t \to \infty$.

43

We have proved that $f_v(t) \geq 0$ if $f_v(0) \geq 0$, and $f_v(t) \leq \sum_{v \in \mathcal{V}} f_v(t) = \sum_{v \in \mathcal{V}} f_v(0)$, i.e. $f_v(t)$ is bounded above by a constant, then by (4.20) and (4.21),

$$\sum_{J(0)} \sum_{q=0}^{p(J(0))} c_{h^q}[(e^{At})h^q]_v = \sum_{J(0)} \sum_{q=0}^{p(J(0))} c_{h^q}(h_v^q + th_v^{q-1} + \dots + \frac{t^q}{q!}h_v^0)$$

is a bounded polynomial, i.e. it must be a constant. So set $t = 0$, we have $\forall v \in \mathcal{V}$

$$\sum_{J(0)} \sum_{q=0}^{p(J(0))} c_{h^q}[(e^{At})h^q]_v = \sum_{J(0)} \sum_{q=0}^{p(J(0))} c_{h^q}h_v^q,$$

and

$$\sum_{J(0)} \sum_{q=1}^{p(J(0))} c_{h^q}(th_v^{q-1} + \dots + \frac{t^q}{q!}h_v^0) = 0.$$

Since the eigenvector and generalized eigenvectors corresponding to distinct Jordan blocks are linearly independent, we have for any Jordan block corresponding to the eigenvalue 0,

$$\sum_{q=1}^{p(J(0))} c_{h^q}(th^{q-1} + \dots + \frac{t^q}{q!}h^0) = \mathbf{0},$$

i.e., if $p = p(J(0))$,

$$c_{h^1}(th^0) + c_{h^2}[th^1 + \frac{t^2}{2!}h^0] + \dots + c_{h^p}[th^{p-1} + \dots + \frac{t^p}{p!}h^0] = \mathbf{0},$$

where $h^0, ..., h^p$ are linearly independent. Hence $c_{h^p} = 0$, which implies that $c_{h^{p-1}} = 0$, and so on...; Hence

$$c_{h^q} \equiv 0 \ \ \forall 1 \leq q \leq p.$$

We obtain the following theorem:

Theorem 4.5: If $(\mathcal{V}, E)$ is a finite directed graph with no loops, and

$$\frac{d}{dt}f_v(t) = \sum_{(v',v) \in E} W(v', v)f_{v'}(t) - [\sum_{(v,v'') \in E} W(v, v'')]f_v(t),$$

where $f_v(0) \geq 0 \ \forall v \in \mathcal{V}$, and $f_v(0) = \sum_{\lambda} \sum_{J(\lambda)} \sum_{q=0}^{p(J(\lambda))} c_{h^q}h_v^q$, then for $t \geq 0$,

$$f_v(t) = \sum_{h^0} c_{h^0}h_v^0 + O(e^{-\epsilon t}),$$

44

where h^0 ranges over the eigenvectors belonging to the Jordan blocks $J(0)$ belonging to the eigenvalue 0 of the matrix A, $\epsilon = \frac{1}{2} \min_{\lambda} |\mathrm{Re}\lambda|$, λ ranges over all eigenvalues of A with $\mathrm{Re}\lambda < 0$.

Note: It follows that $\sum_{h^0} c_{h^0} h_v^0 \geq 0$ for all $v \in \mathcal{V}$. Although A must have at least one nonnegative eigenvector h^0 belonging to the eigenvalue 0 (by Perron-Frobenius theorem), we do not know if all these eigenvectors are nonnegative.

Chapter 5

Construction of the Markov Process

5.1 Products of Countably Many Measure Spaces

A key step in the construction of a Markov process is to find a probability space $(\mathcal{T}_\gamma, \mathcal{M}_\gamma, P_\gamma)$ on which a sufficiently rich family of independent random variables can be defined. Infinite products of probability spaces are the natural chocies for $(\mathcal{T}_\gamma, \mathcal{M}_\gamma, P_\gamma)$.

Let $(\mathcal{T}_\gamma, \mathcal{M}_\gamma, P_\gamma)$ denote a measure space for each γ contained in an index set Γ, where $\Gamma = \mathbb{N} = \{1, 2, 3, ...\}$, and $P_\gamma(\mathcal{T}_\gamma) = 1$ for every $\gamma \in \Gamma$. Let $\mathcal{T} = \prod_{\gamma \in \Gamma} \mathcal{T}_\gamma$, and for $\underline{t} \in \mathcal{T}$ let $\underline{t} = (t_\gamma)$, where t_γ denotes the value of $\underline{t}$ at γ. Let Ω, with or without a subscript, be reserved for finite subsets of Γ. For any $\Omega = \{\gamma_1, \gamma_2, ..., \gamma_n\} \subset \Gamma$, let ε_Ω be the collection of all sets of the form

$$A_\Omega = A_{\gamma_1} \times A_{\gamma_2} \times \cdots \times A_{\gamma_n},$$

where $A_{\gamma_j} \in \mathcal{M}_{\gamma_j} \; \forall j$. If $\Delta \subset \Gamma$, let $\Delta' = \Gamma \setminus \Delta$, and $\mathcal{T}_\Delta = \prod_{\gamma \in \Delta} \mathcal{T}_\gamma$. In particular, $\mathcal{T} = \mathcal{T}_\Gamma$. Let $\mathcal{N}_\Delta$ be the smallest algebra of subsets of $\mathcal{T}_\Delta$ that contains all sets of the form $A_\Omega \times \mathcal{T}_{\Delta \cap \Omega'}$ (the product is not ordered), where Ω runs through all finite subsets of Δ and A_Ω through all sets in ε_Ω. Let $\mathcal{M}_\Delta$ be the σ-algebra generated by $\mathcal{N}_\Delta$, and write $\mathcal{N}$ for $\mathcal{N}_\Gamma$, $\mathcal{M}$ for $\mathcal{M}_\Gamma$.

Our goal in this section is to construct a unique countably additive measure $\overline{P}$ on $\mathcal{M}$ such that

$$(1) \quad \overline{P}(\mathcal{T}) = 1,$$

$$(2) \quad \overline{P}((A_{\gamma_1} \times \cdots \times A_{\gamma_n}) \times \mathcal{T}_{\{\gamma_1,\ldots,\gamma_n\}'}) = P_{\gamma_1}(A_{\gamma_1}) \cdots P_{\gamma_n}(A_{\gamma_n}), \qquad (5.1)$$

where $A_{\gamma_j} \in \mathcal{M}_{\gamma_j} \; \forall j$.

We will use the following 5 lemmas to prove the above statement in this section. The proof is adapted from ([12]).

Lemma 5.1: Let $\Omega = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$ be any finite nonempty subset of Γ, then there is a unique measure P_Ω on $\mathcal{M}_\Omega$ such that

$$P_\Omega(\prod_{j=1}^{n} A_{\gamma_j}) = \prod_{j=1}^{n} P_{\gamma_j}(A_{\gamma_j}) \qquad (5.2)$$

for all $\prod_{j=1}^{n} A_{\gamma_j} \in \varepsilon_\Omega$.

Lemma 5.2: If $\Omega_1 \cap \Omega_2 = \emptyset$ and $B_{\Omega_j} \in \mathcal{M}_{\Omega_j} (j = 1, 2)$, then

$$P_{\Omega_1 \cup \Omega_2}(B_{\Omega_1} \times B_{\Omega_2}) = P_{\Omega_1}(B_{\Omega_1}) \cdot P_{\Omega_2}(B_{\Omega_2}). \qquad (5.3)$$

Lemma 5.3: Let $\mathcal{T}$ be an arbitrary set and $\mathcal{N}$ an algebra of subsets of $\mathcal{T}$. Let P be a set function on $\mathcal{N}$ satisfying the following conditions:

(1) $0 \leq P(A) \leq \infty$ for all $A \in \mathcal{N}$;

(2) $P(A \cup B) = P(A) + P(B)$ for $A, B \in \mathcal{N}$ and $A \cap B = \emptyset$;

(3) if $A_1, A_2, \ldots \in \mathcal{N}$, $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$, and $\cap_{n=1}^{\infty} A_n = \emptyset$, then

$$\lim_{n \to \infty} P(A_n) = 0.$$

For $S \subset T$, let

$$\overline{P}(S) = \inf\{\sum_{n=1}^{\infty} P(A_n) : S \subset \cup_{n=1}^{\infty} A_n, \text{ and } A_1, ..., A_n, ... \in \mathcal{N}\}, \qquad (5.4)$$

then P can be extended to the countably additive measure $\overline{P}$ defined on the collection of all $\overline{P}$−measurable subsets of T (A set $A \in T$ is $\overline{P}$ measurable if for all set $E \in T$, $\overline{P}(E) = \overline{P}(E \cap A) + \overline{P}(E \setminus A)$), which is a σ−algebra of subsets of T that contains $\mathcal{N}$. This lemma is given as (10.37) in ([12]).

Lemma 5.4: Let T be an arbitrary set and $\mathcal{N}$ an algebra of subsets of T. If the family $\mathcal{F}$ of subsets of T contains $\mathcal{N}$ and satisfies the following two conditions:

(1) If $E_n \in \mathcal{F}$ and $E_n \subset E_{n+1}$ for $n = 1, 2, 3, ...$, then $\cup_{n=1}^{\infty} E_n \in \mathcal{F}$;

(2) If $F_n \in \mathcal{F}$ and $F_n \supset F_{n+1}$ for $n = 1, 2, 3, ...$, then $\cap_{n=1}^{\infty} F_n \in \mathcal{F}$.

then $\mathcal{F}$ contains the σ−algebra $\mathcal{M}$ generated by $\mathcal{N}$. This lemma is given as (21.6) in ([12]).

Lemma 5.5: Let $(T, \mathcal{M}, \overline{P})$ be a measure space. If $\{E_n\}_{n=1}^{\infty}$ is a sequence of sets in $\mathcal{M}$ such that $E_1 \subset E_2 \subset \cdots \subset E_n \subset \cdots$, then

$$P(\cup_{n=1}^{\infty} E_n) = \lim_{n \to \infty} P(E_n);$$

If $\{F_n\}_{n=1}^{\infty}$ is a sequence of sets in $\mathcal{M}$ such that $P(F_1) < \infty$ and $F_1 \supset F_2 \supset \cdots \supset F_n \supset \cdots$, then

$$P(\cap_{n=1}^{\infty} F_n) = \lim_{n \to \infty} P(F_n).$$

This lemma is given as (10.15) in ([12]).

We begin with the following theorem.

Theorem 5.1: There is a unique finitely additive measure P on the algebra of sets $\mathcal{N}$ such that

$$P(A_\Omega \times T_{\Omega'}) = P_\Omega(A_\Omega) \qquad (5.5)$$

48

for all Ω and all $A_\Omega \in \mathcal{M}_\Omega$.

Proof Let (5.5) define P. First we show that P is well defined. Suppose

$$A_{\Omega_1} \times \mathcal{T}_{\Omega_1'} = A_{\Omega_2} \times \mathcal{T}_{\Omega_2'}.$$

Let $\Omega_3 = \Omega_1 \cup \Omega_2$. Then (5.5) can be rewritten as

$$A_{\Omega_1} \times \mathcal{T}_{\Omega_3 \cap \Omega_1'} \times \mathcal{T}_{\Omega_3'} = A_{\Omega_2} \times \mathcal{T}_{\Omega_3 \cap \Omega_2'} \times \mathcal{T}_{\Omega_3'}. \tag{5.6}$$

Since the sets $A_{\Omega_1} \times \mathcal{T}_{\Omega_3 \cap \Omega_1'}$ and $A_{\Omega_2} \times \mathcal{T}_{\Omega_3 \cap \Omega_2'}$ are in $\mathcal{M}_{\Omega_3}$, (5.6) shows that they are equal. By Lemma 5.2, we have

$$P_{\Omega_3}(A_{\Omega_1} \times \mathcal{T}_{\Omega_3 \cap \Omega_1'}) = P_{\Omega_1}(A_{\Omega_1}) \cdot 1 = P_{\Omega_1}(A_{\Omega_1}),$$

and

$$P_{\Omega_3}(A_{\Omega_2} \times \mathcal{T}_{\Omega_3 \cap \Omega_2'}) = P_{\Omega_2}(A_{\Omega_2}) \cdot 1 = P_{\Omega_2}(A_{\Omega_2}).$$

Hence

$$P_{\Omega_1}(A_{\Omega_1}) = P_{\Omega_2}(A_{\Omega_2}),$$

i.e. P is well defined.

We claim that for any set $A \in \mathcal{N}$, there exists an Ω and a set $B_\Omega \in \mathcal{N}_\Omega$, such that

$$A = B_\Omega \times \mathcal{T}_{\Omega'}.$$

The proof is given as follows.

Let $\mathcal{C} = \{A \in \mathcal{N} : A = B_\Omega \times \mathcal{T}_{\Omega'}, B_\Omega \in \mathcal{N}_\Omega, \Omega \subset \Gamma\}$, then $\mathcal{C} \subset \mathcal{N}$. It is clear that $\emptyset \in \mathcal{C}$, $\mathcal{T} \in \mathcal{C}$, and for any $\Omega \subset \Gamma$, the subset of $\mathcal{T}_\Omega$ of the form $A_\Omega \times \mathcal{T}_{\Omega'}$ where $A_\Omega \in \varepsilon_\Omega$ is in $\mathcal{C}$. To prove our claim, it suffices to show that $\mathcal{C}$ is an algebra, then $\mathcal{C} \supset \mathcal{N}$ by the definition of $\mathcal{N}$, and then we obtain $\mathcal{C} = \mathcal{N}$.

$\mathcal{C}$ is closed under finite union: let $A_1 \in \mathcal{C}$ and $A_2 \in \mathcal{C}$, where $A_1 = B_{\Omega_1} \times \mathcal{T}_{\Omega_1'}, B_{\Omega_1} \in \mathcal{N}_{\Omega_1}$, and $A_2 = B_{\Omega_2} \times \mathcal{T}_{\Omega_2'}, B_{\Omega_2} \in \mathcal{N}_{\Omega_2}$. Let $\Omega_3 = \Omega_1 \cup \Omega_2$, then

$$
\begin{aligned}
A_1 \cup A_2 &= B_{\Omega_1} \times \mathcal{T}_{\Omega_1'} \cup B_{\Omega_2} \times \mathcal{T}_{\Omega_2'} \\
&= B_{\Omega_1} \times \mathcal{T}_{\Omega_3 \cap \Omega_1'} \times T_{\Omega_3'} \cup B_{\Omega_2} \times \mathcal{T}_{\Omega_3 \cap \Omega_2'} \times \mathcal{T}_{\Omega_3'} \\
&= (B_{\Omega_1} \times \mathcal{T}_{\Omega_3 \cap \Omega_1'} \cup B_{\Omega_2} \times \mathcal{T}_{\Omega_3 \cap \Omega_2'}) \times \mathcal{T}_{\Omega_3'}.
\end{aligned}
$$

Since $\mathcal{N}_{\Omega_j} \times \mathcal{T}_{\Omega_3 \cap \Omega_j'}$ is the smallest algebra of subsets of $\mathcal{T}_{\Omega_3}$ that contains all sets $B_{\Omega_j} \times \mathcal{T}_{\Omega_3 \cap \Omega_j'}$, where $B_{\Omega_j} \in \varepsilon_{\Omega_j}$, we have $\mathcal{N}_{\Omega_j} \times \mathcal{T}_{\Omega_3 \cap \Omega_j'} \subset \mathcal{N}_{\Omega_3}$. That is, $B_{\Omega_1} \times \mathcal{T}_{\Omega_3 \cap \Omega_1'} \in \mathcal{N}_{\Omega_3}$ and $B_{\Omega_2} \times \mathcal{T}_{\Omega_3 \cap \Omega_2'} \in \mathcal{N}_{\Omega_3}$, and their union is also in $\mathcal{N}_{\Omega_3}$, hence $A_1 \cup A_2 \in \mathcal{C}$.

And $\mathcal{C}$ is closed under complement: let $A \in \mathcal{C}$, then $A = B_{\Omega} \times \mathcal{T}_{\Omega'}$ for some $\Omega \subset \Gamma$, where $B_{\Omega} \in \mathcal{N}_{\Omega}$. We have

$$
A' = (B_{\Omega} \times \mathcal{T}_{\Omega'})' = B_{\Omega}' \times \mathcal{T}_{\Omega'}.
$$

Since $B_{\Omega} \in \mathcal{N}_{\Omega}$, then $B_{\Omega}' \in \mathcal{N}_{\Omega}$, i.e. $A' \in \mathcal{C}$. This proves our claim.

Let A_1 and A_2 be two disjoint sets in $\mathcal{N}$, where $A_1 = B_{\Omega_1} \times \mathcal{T}_{\Omega_1'}, B_{\Omega_1} \in \mathcal{N}_{\Omega_1}$, and $A_2 = B_{\Omega_2} \times \mathcal{T}_{\Omega_2'}, B_{\Omega_2} \in \mathcal{N}_{\Omega_2}$. Let $\Omega_3 = \Omega_1 \cup \Omega_2$, then write

$$
A_j = B_{\Omega_3}^{(j)} \times \mathcal{T}_{\Omega_3'} \quad (j = 1, 2),
$$

where $B_{\Omega_3}^{(j)} \in \mathcal{N}_{\Omega_3}$. And since $B_{\Omega_3}^{(1)} \cap B_{\Omega_3}^{(2)} = \emptyset$, then using (5.5) and the additivity of P_{Ω_3} from lemma 5.1, we have

$$
\begin{aligned}
P(A_1 \cup A_2) &= P_{\Omega_3}(B_{\Omega_3}^{(1)} \cup B_{\Omega_3}^{(2)}) \\
&= P_{\Omega_3}(B_{\Omega_3}^{(1)}) + P_{\Omega_3}(B_{\Omega_3}^{(2)}) \\
&= P(A_1) + P(A_2),
\end{aligned}
$$

i.e., P is finitely additive on $\mathcal{N}$. $\square$

Theorem 5.2: The finitely additive measure P on $\mathcal{N}$ admits a unique extension over $\mathcal{M}$ that is countably additive.

Proof *Existence*: By lemma 5.3, to prove that P has some countably additive extension over $\mathcal{M}$, we need only show that

$$\lim_{n \to \infty} P(F_n) = 0$$

for every sequence $(F_n)_{n=1}^{\infty}$ such that $F_n \in \mathcal{N}$, $F_1 \supset F_2 \supset \cdots \supset F_n \supset \cdots$, and $\cap_{n=1}^{\infty} F_n = \emptyset$. Each F_n has the form $B_{\Omega_n} \times \mathcal{T}_{\Omega_n'}$, where $B_{\Omega_n} \in \mathcal{N}_{\Omega_n}$.

Let $k_n = \max(\Omega_n)$. Without loss of generality, we can assume that $\Omega_n = \{1, 2, ..., k_n\}$ and that $k_1 < k_2 < k_3 < \cdots$. Define the sequence of sets $(E_m)_{m=1}^{\infty}$ as follows:

$$E_m = \begin{cases} \mathcal{T} & \text{if } 1 \le m < k_1, \\ F_n & \text{if } k_m \le m < k_{n+1}. \end{cases}$$

Then we have $\cap_{m=1}^{\infty} E_m = \cap_{n=1}^{\infty} F_n = \emptyset$, and $\lim_{m \to \infty} P(E_m) = \lim_{n \to \infty} P(F_n)$. Hence to prove $\lim_{n \to \infty} P(F_n) = 0$, we need to show $\lim_{m \to \infty} P(E_m) = 0$. Let $\Theta_m = \{1, 2, ..., m\}$ for each m, then each E_m has the form $B_m \times \mathcal{T}_{\Theta_m'}$, where $B_m \in \mathcal{N}_{\Theta_m}$.

From lemma 5.2, we have $P_{\Theta_{m+1}} = P_{\Theta_{m+1}} \times P_m$ for all m, and a set in $\mathcal{N}_{\Theta_{m+1}}$ is $\mathcal{M}_{\Theta_m} \times \mathcal{M}_{m+1}$— measurable. By the Fubini Theorem and theorem 5.1, we obtain

$$P(E_m) = P_{\Theta_m}(B_m)$$

$$= \int_{\mathcal{T}_{\Theta_m}} 1_{B_m}(\underline{t}) d_{P_{\Theta_m}}(\mathbf{t})$$

$$= \int_{\mathcal{T}_{\Theta_{m-1}}} \int_{\mathcal{T}_m} 1_{B_m}(\mathbf{t}*, t_m) d_{P_m}(t_m) d_{P_{\Theta_{m-1}}}(\mathbf{t}*), \qquad (5.7)$$

where $\mathbf{t}*$ denotes a generic element of $\mathcal{T}_{\Theta_{m-1}}$. By the Fubini Theorem, the inner integral in (5.7) is $\mathcal{M}_{\Theta_{m-1}}$— measurable, then we can apply the Fubini Theorem and theorem 5.1 again. After doing this $m - 1$ times, we have

$$P(E_m) = \int_{\mathcal{T}_1} \int_{\mathcal{T}_2} \cdots \int_{\mathcal{T}_m} 1_{B_m}(t_1, t_2, ..., t_m) d_{P_m}(t_m) \cdots d_{P_2}(t_2) d_{P_1}(t_1).$$

We will prove by contradiction: Assume that $\lim_{m \to \infty} P(E_m) \ne 0$. For $s_1 \in \mathcal{T}_1$, let

$$f_{1,m}(s_1) = \int_{\mathcal{T}_2} \cdots \int_{\mathcal{T}_m} 1_{B_m}(s_1, t_2, ..., t_m) d_{P_m}(t_m) \cdots d_{P_2}(t_2)$$

for $m = 2, 3, ...$, then

$$P(E_m) = \int_{\mathcal{T}_1} f_{1,m}(t_1)dp_1(t_1),$$

and $f_{1,m}(\mathcal{T}_1) \subset [0, 1]$. By our assumption, we know that $\lim_{m \to \infty} f_{1,m}(t_1) = 0$ cannot hold true everywhere on $\mathcal{T}_1$; otherwise, the dominated convergence theorem would imply that

$$\lim_{m \to \infty} P(E_m) = \lim_{m \to \infty} \int_{\mathcal{T}_1} f_{1,m}(t_1)dp_1(t_1) = \int_{\mathcal{T}_1} \lim_{m \to \infty} f_{1,m}(t_1)dp_1(t_1) = 0.$$

Hence there exist a point $a_1 \in \mathcal{T}_1$ such that $\lim_{m \to \infty} f_{1,m}(a_1) \neq 0$.

Next define

$$f_{2,m}(s_2) = \int_{\mathcal{T}_3} \cdots \int_{\mathcal{T}_m} 1_{B_m}(a_1, s_2, t_3, ..., t_m)dp_m(t_m) \cdots dp_3(t_3)$$

for $m = 3, ...$. Similarly, if it is true that $\lim_{m \to \infty} f_{2,m}(t_2) = 0$ for all $s_2 \in \mathcal{T}_2$, then we would also have

$$\lim_{m \to \infty} \int_{\mathcal{T}_2} f_{2,m}(t_2)dp_2(t_2) = \lim_{m \to \infty} f_{1,m}(a_1) = 0.$$

Hence there is a point $a_2 \in \mathcal{T}_2$ such that $\lim_{m \to \infty} f_{2,m}(a_2) \neq 0$.

In this way we construct a sequence of points $\mathbf{a} = (a_1, a_2, ..., a_n, ...)$ in $\mathcal{T}$ with the following property: for every $n \in N$, the sequence of numbers

$$f_{n,m}(a_n) = \int_{\mathcal{T}_{n+1}} \int_{\mathcal{T}_{n+2}} \cdots \int_{\mathcal{T}_m} 1_{B_m}(a_1, a_2, ..., a_n, t_{n+1}, ..., t_m)dp_m(t_m) \cdots dp_{n+1}(t_{n+1})$$

$$(5.8)$$

for $m > n$, does not converge to 0 as $m \to 0$. Therefore the integrand in (5.8) cannot be identically equal to 0 for all large m. So

$$(a_1, a_2, ..., a_n, s_{n+1}, ..., s_m) \in B_m$$

for appropriate $s_j \in \mathcal{T}_j$, and for any large m. Since $E_m = B_m \times \mathcal{T}_{\Theta'_m}$, we can choose a point $\mathbf{s}^{(\mathbf{m},\mathbf{n})} \in \mathcal{T}_{\Theta'_m}$ such that

$$(a_1, a_2, ..., a_n) \times \mathbf{s}^{(\mathbf{m},\mathbf{n})} \in E_m.$$

52

Since $m > n$, we have $E_m \subset E_n$, hence

$$(a_1, a_2, ..., a_n) \times \mathbf{s^{(m,n)}} \in E_n,$$

which implies

$$(a_1, a_2, ..., a_n) \times \mathcal{T}_{\Theta'_n} \subset E_n,$$

since $E_n = B_n \times \mathcal{T}_{\Theta'_n}$. In particular, $\mathbf{a} \in E_n \ \forall n \geq 1$.

Since n is an arbitrary positive integer, we have

$$\mathbf{a} \in \cap_{n=1}^{\infty} E_n.$$

This contradicts $\cap_{n=1}^{\infty} E_n = \emptyset$.

Thus now we have a countably additive measure $\overline{P}$ on the $\sigma-$algebra $\mathcal{M}$ of subsets of $\mathcal{T}$, which is extended from the finitely additive measure P on $\mathcal{N}$ according to Lemma 5.3.

Uniqueness: Suppose there exists two different countably additive measures extended from P, denoted by $\overline{P}$ and $\overline{P'}$, on $\mathcal{M}$. Define

$$\mathcal{F} = \{F \in \mathcal{M} : \overline{P}(F) = \overline{P'}(F)\}.$$

We claim that $\mathcal{F}$ is the $\sigma-$algebra generated by $\mathcal{N}$, i.e., $\mathcal{F} = \mathcal{M}$. Hence $\overline{P}$ is unique on $\mathcal{M}$.

It is obvious that $\mathcal{N} \subset \mathcal{F}$ since $\overline{P}$ and $\overline{P'}$ on $\mathcal{F}$ are extended from the same measure P on $\mathcal{N}$. If we can prove that $\mathcal{F}$ satisfies the two conditions in lemma 5.4, then $\mathcal{F}$ is a $\sigma-$algebra containing $\mathcal{N}$, i.e., $\mathcal{F}$ contains $\mathcal{M}$. And since $\mathcal{F} \subset \mathcal{M}$, we obtain $\mathcal{F} = \mathcal{M}$.

Let $\{E_n\}_{n=1}^{\infty}$ be the sequence of elements in $\mathcal{F}$ such that $E_1 \subset E_2 \subset \cdots \subset E_n \subset \cdots$, and $\{F_n\}_{n=1}^{\infty}$ be the sequence of elements in $\mathcal{F}$ such that $F_1 \supset F_2 \supset \cdots \supset F_n \supset \cdots$,

then $E_n \in \mathcal{M}$, $F_n \in \mathcal{M}$, $\overline{P}(E_n) = \overline{P'}(E_n)$, and $\overline{P}(F_n) = \overline{P'}(F_n)$ for all n. Let $E = \cup_{n=1}^{\infty} E_n$, $F = \cap_{n=1}^{\infty} F_n$, since $\overline{P}(\mathcal{T}) = \overline{P'}(\mathcal{T}) = 1 < \infty$, by lemma 5.5, we have

$$\overline{P}(E) = \lim_{n \to \infty} \overline{P}(E_n) = \lim_{n \to \infty} \overline{P'}(E_n) = \overline{P'}(E),$$

and

$$\overline{P}(F) = \lim_{n \to \infty} \overline{P}(F_n) = \lim_{n \to \infty} \overline{P'}(F_n) = \overline{P'}(F).$$

Hence $E \in \mathcal{F}$ and $F \in \mathcal{F}$, i.e. $\mathcal{F}$ satisfies the two conditions in lemma 5.4. $\square$

Theorem 5.3: There exists a countable family of independent identically distributed (i.i.d) random variables on the probability space $(\mathcal{T}, \mathcal{M}, \overline{P})$, where $(\mathcal{T}, \mathcal{M}, \overline{P})$ is the product of countably many copies of $(\mathbb{R}, \mathcal{M}_{\mathbb{R}}, P_{\mathbb{R}})$, where $\mathcal{M}_{\mathbb{R}}$ is a $\sigma-$algebra on $\mathbb{R}$ and $P_{\mathbb{R}} : \mathcal{M}_{\mathbb{R}} \longrightarrow [0,1]$ is a probability measure describing the distribution of each of these random variables.

Proof: For $n \geq 1$, define

$$\pi_n : \mathcal{T} \to \mathbb{R},$$

i.e.

$$\pi_n(t_1, t_2, ...) = t_n.$$

Let $J = \{j_1, ..., j_m\}$ be an arbitrary subset of $\mathbb{N} = \{1, 2, 3, ...\}$, where $j_i < j_k$ if $i < k$, and $I_j \in \mathcal{M}_{\mathbb{R}}$ where $j \in J$. Observe that $\pi_j^{-1}(I_j)$ is the collection of all $(t_n)_{n=1}^{\infty}$ such that $t_j \in I_j$, i.e.

$$\pi_j^{-1}(I_j) = \mathbb{R} \times^1 \cdots \times^{j-2} \mathbb{R} \times^{j-1} I_j \times^j \mathcal{T}.$$

Thus

$$\bigcap_{k=1}^{m} \pi_{j_k}^{-1}(I_{j_k}) = \mathbb{R} \times^1 \cdots \times^{j_1-2} \mathbb{R} \times^{j_1-1} I_{j_1} \times^{j_1} \mathbb{R} \times \cdots \times \mathbb{R} \times^{j_m-1} I_{j_m} \times^{j_m} \mathcal{T}.$$

Also $\overline{P}(\pi_j^{-1}(I_j)) = P_{\mathbb{R}}(I_{j_k})$ showing that π_j has the same distribution $P_{\mathbb{R}}$ $\forall j \geq 1$.

By (5.1) we have

$$\overline{P}(\bigcap_{k=1}^{m} \pi_{j_k}^{-1}(I_{j_k})) = P_{\mathbb{R}}(I_{j_1}) \cdots P_{\mathbb{R}}(I_{j_m}).$$

And since $\forall j_k \in J$,

$$\overline{P}(\pi_{j_k}^{-1}(I_{j_k})) = P_{\mathbb{R}}(I_{j_k}),$$

then

$$\overline{P}(\bigcap_{k=1}^{m} \pi_{j_k}^{-1}(I_{j_k})) = \prod_{k=1}^{m} \overline{P}(\pi_{j_k}^{-1}(I_{j_k})), \tag{5.9}$$

i.e. $\{\pi_n\}_{n=1}^{\infty}$ is a family of i.i.d random variables on $(\mathcal{T}, \mathcal{M}, \overline{P})$. $\square$

5.2 Facts about the Exponentially Distributed Random Variables

Before constructing the Markov process, we first introduce three facts about the exponentially distributed random variables. If $\alpha > 0$, then the random variable V is said to be exponentially distributed with parameter α if $P(V > t) = e^{-\alpha t} \ \forall t > 0$.

Property 1: Let V be exponentially distributed with unit mean, then $\frac{V}{\alpha}$ is exponentially distributed with parameter α.

Proof: We have

$$P(V > t) = e^{-t},$$

hence

$$P(\frac{V}{\alpha} > t) = P(V > \alpha t) = e^{-\alpha t},$$

i.e. $\frac{V}{\alpha}$ is exponentially distributed with parameter α. $\square$

Property 2: Let $E(\alpha)$ be exponentially distributed with parameter α, then $E(\alpha)$ has the forgetfulness property: for $s, t > 0$,

$$P(E(\alpha) > t + s \mid E(\alpha) > s) = P(E(\alpha) > t) = e^{-\alpha t}.$$

Proof: We calculate

$$
\begin{aligned}
P(E(\alpha) > t + s \mid E(\alpha) > s) &= \frac{P(E(\alpha) > t + s, E(\alpha) > s)}{P(E(\alpha) > s)} \\
&= \frac{P(E(\alpha) > t + s)}{P(E(\alpha) > s)} \\
&= \frac{e^{-\alpha(t+s)}}{e^{-\alpha s}} \\
&= e^{-\alpha t} \\
&= P(E(\alpha) > t). \ \square
\end{aligned}
$$

Property 3: Let $E(\alpha)$ and $E(\beta)$ be independent with exponential distributions with parameter α, β respectively, where $\alpha > 0$, $\beta > 0$. Denote the minimum of two numbers x, y by $x \wedge y$, then $E(\alpha) \wedge E(\beta)$ is exponentially distributed with parameter $\alpha + \beta$ for $t > 0$, i.e.

$$
P(E(\alpha) \wedge E(\beta) > t) = e^{-(\alpha+\beta)t}.
$$

Proof: Since $E(\alpha)$ and $E(\beta)$ are independent, we have

$$
\begin{aligned}
P(E(\alpha) \wedge E(\beta) > t) &= P(E(\alpha) > t, E(\beta) > t) \\
&= e^{-\alpha t} e^{-\beta t} \\
&= e^{-(\alpha+\beta)t}. \ \square
\end{aligned}
$$

Given the state of a stochastic process is $\mathcal{X}'$ at time t, we claim that the holding time of state $\mathcal{X}'$ is exponentially distributed with parameter $W(\mathcal{X}', \mathcal{X})$ under the assumption of (4.6): Let $E(W(\mathcal{X}', \mathcal{X}))$, which is exponentially distributed with parameter $W(\mathcal{X}', \mathcal{X})$, denote the holding time of state $\mathcal{X}'$, then by Property 2, the probability that a reaction r, which connects the states $\mathcal{X}'$ and $\mathcal{X}$, would occur once in $(t, t + dt)$ is

$$
P(E(W(\mathcal{X}', \mathcal{X})) < t+dt \mid E(W(\mathcal{X}', \mathcal{X})) \geq t) = 1 - e^{-W(\mathcal{X}', \mathcal{X})dt} = W(\mathcal{X}', \mathcal{X})dt + o(dt).
$$

5.3 Construction of the Markov Process on an Infinite Product Space

We want to construct a Markov process $\{X(t), t \geq 0\}$ on the probability space $(\mathcal{T}, \mathcal{M}, P)$ which describes the time evolution of a random walk on $\mathcal{V}$, where $(\mathcal{V}, E)$ is a directed graph with no loops, $|\mathcal{V}| < \infty$, and $W : E \longrightarrow (0, \infty)$ giving the probability per unit time of a transition along a given edge. But we will describe the construction in the case where $(\mathcal{V}, E)$ is a State graph of a labeled Petri net and an interger k. Let $\mathcal{T} = (\prod\limits_{m=0}^{\infty}(0, 1]) \times (\prod\limits_{n=0}^{\infty}(0, \infty))$, and P be defined to be $\overline{P}$ as in section 5.1, where $(0, 1]$ has the Lebesgue measure, and $(0, \infty)$ has the measure $e^{-v}dv$. The construction is based on a sequence of i.i.d uniformly distributed random variables $\{U_m\}$ defined on $(0, 1]$, and a sequence of i.i.d exponentially distributed random variables $\{V_n\}$ with unit mean defined on $(0, \infty)$. The existence of U_m and V_n which are all independent has been verified by Theorem 5.3. An element $\underline{t} \in \mathcal{T}$ has the form $(\underline{u}, \underline{v})$, where $\underline{u} = (u_0, u_1, ...)$, $\underline{v} = (v_0, v_1, ...)$, $U_m(\underline{t}) = u_m$, and $V_n(\underline{t}) = v_n$.

The usual approach to construct a continuous time Markov process $\{X(t), t \geq 0\}$ with a countable state space $\mathcal{V}$ follows 2 steps: first, we construct a discrete parameter Markov chain $\{X_n, n = 0, 1, 2, ...\}$ with the state space $\mathcal{V}$ to govern movements through the state space, and then construct the Markov process $\{X(t)\}$ from the Markov chain $\{X_n\}$ and a supply of independent exponentially distributed random variables with unit mean which control how rapidly these movements take place ([18]).

First we will construct the Markov chain $\{X_n, n = 0, 1, 2, ...\}$. We may think of this Markov chain as a discrete time random walk on the directed graph $G = (\mathcal{V}, E)$, where transitions are controlled by the oriented edges in E. Let the set R be linearly ordered by $<$. For each state $\mathcal{X} \in \mathcal{V}$, define

$$R_{\mathcal{X}} = \{r \in R \mid \mathcal{X} - \mathbf{n}_r \in \mathcal{V}\}$$

Then $R_{\mathcal{X}}$ becomes linearly ordered. Define $\beta_{\mathcal{X}} : R_{\mathcal{X}} \longrightarrow (0, 1]$:

$$\beta_{\mathcal{X}}(r) = \sum_{r' < r, \, r' \in R_{\mathcal{X}}} \frac{W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r'})}{\alpha(\mathcal{X})},$$

where $\alpha(\mathcal{X}) = \sum_{r \in R_{\mathcal{X}}} W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r)$. Also let the state space $\mathcal{V}$ be enumerated, i.e. $\mathcal{V} = \{\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, ...\}$.

We start with initial distribution $\{f_{\mathcal{X}}(0)\}_{\mathcal{X} \in \mathcal{V}}$, where $f_{\mathcal{X}}(0) \geq 0$, $\sum_{\mathcal{X} \in \mathcal{V}} f_{\mathcal{X}}(0) = 1$.

Define the initial state

$$X_0(\underline{t}) = \begin{cases} \mathcal{X}_1 & \text{if } 0 < U_0(\underline{t}) \leq f_{\mathcal{X}_1}(0), \\ \mathcal{X}_2 & \text{if } f_{\mathcal{X}_1}(0) < U_0(\underline{t}) \leq \sum_{j=1}^{2} f_{\mathcal{X}_j}(0), \\ \cdots \\ \mathcal{X}_i & \text{if } \sum_{j=1}^{i-1} f_{\mathcal{X}_j}(0) < U_0(\underline{t}) \leq \sum_{j=1}^{i} f_{\mathcal{X}_j}(0), \\ \cdots. \end{cases}$$

Given $\mathcal{X} = X_n(\underline{t})$ for $n \geq 0$, define X_{n+1} as follows:

$$X_{n+1}(\underline{t}) = \mathcal{X} - \mathbf{n}_r,$$

where $r \in R_{\mathcal{X}}$ is such that

$$\beta_{\mathcal{X}}(r) < U_{n+1}(\underline{t}) \leq \beta_{\mathcal{X}}(r) + \frac{W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r)}{\alpha(\mathcal{X})},$$

i.e.

$$\sum_{r' < r, \, r' \in R_{\mathcal{X}}} \frac{W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r)}{\alpha(\mathcal{X})} < U_{n+1}(\underline{t}) \leq \sum_{r' \leq r, \, r' \in R_{\mathcal{X}}} \frac{W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r)}{\alpha(\mathcal{X})}.$$

Observe X_1 is a function of X_0 and U_1 and hence is a function of U_0 and U_1, X_2 is a function of X_1 and U_2 and hence is a function of U_0, U_1 and U_2, and so on, hence in general we have X_n is a function of $U_0, U_1, U_2, ..., U_n$.

We next construct the continuous time Markov process $\{X(t), t \geq 0\}$ based on the discrete Markov chain $\{X_n\}$ and the sequence of independent exponentially distributed random variables $\{V_n\}$.

Let $T_0(\underline{t}) = 0$ be the initial time $\forall \underline{t} \in \mathcal{T}$. For $n \geq 0$, define

$$T_{n+1}(\underline{t}) = T_n(\underline{t}) + \frac{V_n(\underline{t})}{\alpha(X_n(\underline{t}))} = \sum_{i=0}^{n} \frac{V_i(\underline{t})}{\alpha(X_i(\underline{t}))},$$

$$X(t)(\underline{t}) = X_n(\underline{t}), \text{ for } T_n(\underline{t}) \leq t < T_{n+1}(\underline{t}).$$

We construct the holding time of state $\mathcal{X}_n(\underline{t}) = \mathcal{X}$ as $V_n(\underline{t})/\alpha(\mathcal{X})$ because of the following reason: The waiting time of the occurrence of a reaction r from state $\mathcal{X}$ to reach $\mathcal{X}'$ is exponentially distributed with parameter $W(\mathcal{X}, \mathcal{X}')$. The holding time of state $\mathcal{X}$ in fact is the waiting time of the occurrence of the reaction which occurs first, which is exponentially distributed with parameter $\alpha(\mathcal{X}) = \sum_{(\mathcal{X}, \mathcal{X}') \in E} W(\mathcal{X}, \mathcal{X}')$ by Property 3 of the exponentially distributed random variables.

Define $T_\infty(\underline{t}) = \lim_{n \to \infty} T_n(\underline{t})$. We claim that the Markov process we constructed is **Regular** under the assumption that the state space $\mathcal{V}$ is finite, i.e. $P(T_\infty = \infty \mid X_0 = \mathcal{X}_0) = 1$. The proof is given as follows:

Conditional on the sequence of states $\{X_n\}$, T_∞ is a sum of independent exponential random variables with different parameters. By the **Kolmogorov Zero-One Law**, we have the event that T_∞ converges to finite limit has conditional probability 0 or 1 (conditioning on $\{X_n\}$).

By the Proposition 5.2.1 of ([18]):

$$P(T_\infty < \infty \mid X_0 = \mathcal{X}_0, \{X_n\}) = 1 \text{ iff } \sum_{n=1}^{\infty} \frac{1}{\alpha(X_n)} < \infty \quad P - \text{a.s.},$$

we have that $P-$a.s.

$$P(T_\infty < \infty \mid X_0 = \mathcal{X}_0, \{X_n\}) = \begin{cases} 1, & \text{if } \sum_{n=1}^{\infty} \frac{1}{\alpha(X_n)} < \infty \\ 0, & \text{if } \sum_{n=1}^{\infty} \frac{1}{\alpha(X_n)} = \infty. \end{cases}$$

Thus

$$P(T_\infty < \infty \mid X_0 = \mathcal{X}_0, \{X_n\}) = 1_{[\sum_{n=1}^{\infty} \frac{1}{\alpha(X_n)} < \infty]} \quad P - \text{a.s.}$$

59

Taking expectations yields

$$P(T_\infty < \infty \mid X_0 = \mathcal{X}_0) = P(\sum_{n=1}^{\infty} \frac{1}{\alpha(X_n)} < \infty \mid X_0 = \mathcal{X}_0).$$

Since the state space $\mathcal{V}$ is finite, we have $W(\mathcal{X}', \mathcal{X}) < K < \infty$ for some K, $\forall \mathcal{X}', \mathcal{X} \in \mathcal{V}$. Thus $\alpha(X_n) = \sum_{r \in R_{X_n}} W(X_n, X_n - \mathbf{n}_r) < |R|K$, and

$$\sum_{n=1}^{\infty} \frac{1}{\alpha(X_n)} = \sum_{n=1}^{\infty} \frac{1}{\sum_{r \in R_{X_n}} W(X_n, X_n - \mathbf{n}_r)} > \sum_{n=1}^{\infty} \frac{1}{|R|K} = \infty.$$

Hence the Markov process is regular, which indicates that only finitely many reactions will take place in any finite time interval $P-$a.s. if the state space is finite. The claim is proved.

5.4 Properties of the Conditional Probability from the Construction

The process constructed in last section has the Markov property: for any $n > 0$, time points $t_1 < t_2 < ... < t_n < t_{n+1}$ and states $\mathcal{X}_1, ..., \mathcal{X}_{n-1}, \mathcal{X}_n, \mathcal{X}_{n+1}$,

$$P(X(t_{n+1}) = \mathcal{X}_{n+1} \mid X(t_1) = \mathcal{X}_1, ..., X(t_{n-1}) = \mathcal{X}_{n-1}, X(t_n) = \mathcal{X}_n)$$

$$= P(X(t_{n+1}) = \mathcal{X}_{n+1} \mid X(t_n) = \mathcal{X}_n)$$

whenever $P(X(t_1) = \mathcal{X}_1, ..., X(t_{n-1}) = \mathcal{X}_{n-1}, X(t_n) = \mathcal{X}_n) > 0$. This is proved in Section 5.2.1 in ([18]).

We claim that the constructed process has the stationary transition probabilities ([18]), i.e. the transition probability is only dependent on the time interval between two states: for $s, t \geq 0$, and states $\mathcal{X}$ and $\mathcal{X}'$, $P(X(t + s) = \mathcal{X}' \mid X(t) = \mathcal{X})$ is independent of $t \geq 0$. The proof is given as follows:

For $\mathcal{X} \in \mathcal{V}$ and $u \in (0, 1]$, define $J(\mathcal{X}, u) = \mathcal{X} - \mathbf{n}_r$ where r is the unique element of $R_\mathcal{X}$ satisfying

$$\beta_\mathcal{X}(r) < u \le \beta_\mathcal{X}(r) + \frac{W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r)}{\alpha(\mathcal{X})}.$$

Then by the construction $X_{n+1} = J(X_n, U_{n+1})$ for $n \ge 0$. Thus X_{n+1} depends on U_{n+1} explicitly but on $U_0, ..., U_n$ implicitly through the value of X_n.

For $T \ge 0$, $\mathcal{X} \in \mathcal{V}$, and $v \in (0, \infty)$, define $K(T, \mathcal{X}, v) = T + \frac{v}{\alpha(\mathcal{X})}$, then

$$T_{n+1} = K(T_n, X_n, V_n)$$

for $n \ge 0$. Thus T_{n+1} depends on V_n explicitly but on $U_0, ..., U_n, V_0, ..., V_{n-1}$ implicitly through the values of T_n and X_n.

Define $\widetilde{X}_1(\mathcal{X}_0, u_1)$, ..., $\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n)$ recursively by the rule:

$$\widetilde{X}_1(\mathcal{X}_0, u_1) = J(\mathcal{X}_0, u_1),$$
$$\widetilde{X}_{n+1}(\mathcal{X}_0, u_1, ..., u_{n+1}) = J(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), u_{n+1}), \text{ for } n \ge 1.$$

For $t_0 \ge 0$, define $\widetilde{T}_1(\mathcal{X}_0, t_0, v_0)$, ..., $\widetilde{T}_n(\mathcal{X}_0, t_0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1})$ recursively by the rule:

$$\widetilde{T}_1(\mathcal{X}_0, t_0, v_0) = K(t_0, \mathcal{X}_0, v_0),$$
$$\widetilde{T}_{n+1}(\mathcal{X}_0, t_0, u_1, ..., u_n, v_0, ..., v_n) = K(\widetilde{T}_n(\mathcal{X}_0, t_0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}),$$
$$\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), v_n), \text{ for } n \ge 1.$$

We have the following two facts:

$$\widetilde{X}_{n+l}(\mathcal{X}_0, u_1, ..., u_{n+l}) = \widetilde{X}_l(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), u_{n+1}, ..., u_{n+l}),$$

and

$$\widetilde{T}_{n+l}(\mathcal{X}_0, t_0, u_1, ..., u_{n+l-1}, v_0, ..., v_{n+l-1})$$
$$= \widetilde{T}_l(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), \widetilde{T}_n(\mathcal{X}_0, t_0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}),$$
$$u_{n+1}, ..., u_{n+l-1}, v_{n+1}, ..., v_{n+l-1}).$$

61

To prove the first fact, we first check the case when $l = 1$. We have

$$\widetilde{X}_{n+1}(\mathcal{X}_0, u_1, ..., u_{n+1}) = J(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), u_{n+1})$$
$$= \widetilde{X}_1(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), u_{n+1}).$$

Hence the result holds true for $l = 1$. Assuming the result for $l \geq 1$, we next prove the result is true for $l + 1$:

$$\widetilde{X}_{n+l+1}(\mathcal{X}_0, u_1, ..., u_{n+l+1}) = J(\widetilde{X}_{n+l}(\mathcal{X}_0, u_1, ..., u_{n+l}), u_{n+l+1})$$
$$= J(\widetilde{X}_l(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), u_{n+1}, ..., u_{n+l}), u_{n+l+1})$$
$$= \widetilde{X}_{l+1}(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), u_{n+1}, ..., u_{n+l+1}).$$

The second fact holds true when $l = 1$ since

$$\widetilde{T}_{n+1}(\mathcal{X}_0, t_0, u_1, ..., u_n, v_0, ..., v_n)$$
$$= K(\widetilde{T}_n(\mathcal{X}_0, t_0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}), \widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), v_n)$$
$$= \widetilde{T}_1(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), \widetilde{T}_n(\mathcal{X}_0, t_0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}), v_n).$$

Assuming the result is true for $l \geq 1$, then

$$\widetilde{T}_{n+l+1}(\mathcal{X}_0, t_0, u_1, ..., u_{n+l}, v_0, ..., v_{n+l})$$
$$= K(\widetilde{T}_{n+l}(\mathcal{X}_0, t_0, u_1, ..., u_{n+l-1}, v_0, ..., v_{n+l-1}), \widetilde{X}_{n+l}(\mathcal{X}_0, u_1, ..., u_{n+l}), v_{n+l})$$
$$= K(\widetilde{T}_l(\widetilde{X}_n(X_0, u_1, ..., u_n), \widetilde{T}_n(\mathcal{X}_0, t_0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}),$$
$$u_{n+1}..., u_{n+l-1}, v_n, ..., v_{n+l-1}), \widetilde{X}_l(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), u_{n+1}, ..., u_{n+l}), v_{n+l})$$
$$= \widetilde{T}_{l+1}(\widetilde{X}_n(\mathcal{X}_0, u_1, ..., u_n), \widetilde{T}_n(\mathcal{X}_0, t_0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}),$$
$$u_{n+1}..., u_{n+l}, v_n, ..., v_{n+l}).$$

So the result is true for $l + 1$ as well.

Hence we have the formulas: for $n \geq 1$,

$$X_n(\underline{t}) = \widetilde{X}_n(X_0(\underline{t}), U_1(\underline{t}), ..., U_n(\underline{t})),$$
$$T_n(\underline{t}) = \widetilde{T}_n(X_0(\underline{t}), 0, U_1(\underline{t}), ..., U_{n-1}(\underline{t}), V_0(\underline{t}), ..., V_{n-1}(\underline{t})).$$

62

Thus

$$X_{n+l} = \widetilde{X}_{n+l}(X_0, U_1, ..., U_{n+l})$$
$$= \widetilde{X}_l(\widetilde{X}_n(X_0, U_1, ..., U_n), U_{n+1}, ..., U_{n+l})$$
$$= \widetilde{X}_l(X_n, U_{n+1}, ..., U_{n+l}).$$

Also

$$T_{n+l} = \widetilde{T}_{n+l}((X_0, 0, U_1, ..., U_{n+l-1}, V_0, ..., V_{n+l-1}))$$
$$= \widetilde{T}_l(\widetilde{X}_n(X_0, U_1, ..., U_n), \widetilde{T}_n(X_0, 0, U_1, ..., U_{n-1}, V_0, ..., V_{n-1}),$$
$$U_{n+1}, ..., U_{n+l-1}, V_{n+1}, ..., V_{n+l-1})$$
$$= \widetilde{T}_l(X_n, T_n, U_{n+1}, ..., U_{n+l-1}, V_{n+1}, ..., V_{n+l-1}).$$

To verify that the process has the stationary transition probabilities, it suffices to prove that

$$P(X(t+s) = \mathcal{X} \mid X(t) = \mathcal{X}') = P(X(s) = \mathcal{X} \mid X(0) = \mathcal{X}').$$

We have

$$P(X(t) = \mathcal{X}') = \sum_{n=0}^{\infty} \int_{(0,1]^{n+1}} \int_{(0,\infty)^{n+1}} P(X_n = \mathcal{X}', T_n \leq t < T_{n+1}$$
$$\mid U_i = u_i, V_i = v_i, 0 \leq i \leq n)$$
$$e^{-\sum_{i=0}^{n} v_i} \, du_0, ..., dv_n$$
$$= \sum_{n=0}^{\infty} \int_{(0,1]^{n+1}} \int_{(0,\infty)^{n+1}} 1_A(\underline{t}) e^{-\sum_{i=0}^{n} v_i} \, du_0, ..., dv_n,$$

where

$$A = \{\underline{t} \mid \widetilde{X}_n(X_0(u_0), u_1, ..., u_n) = \mathcal{X}',$$
$$\widetilde{T}_n(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}) \leq t,$$
$$t - \widetilde{T}_n(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}) < \frac{v_n}{\alpha(\mathcal{X}')}\},$$

and

$$\widetilde{T_n}(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}) = \frac{v_0}{\alpha(X_0(u_0))} + \sum_{i=1}^{n-1} \frac{v_i}{\alpha(\widetilde{X}_i(X_0(u_0), u_1, ..., u_i))}.$$

Then A can be written as the intersection of two sets:

$$A_1 = \{\underline{t} \mid \widetilde{X}_n(X_0(u_0), u_1, ..., u_n) = \mathcal{X}', \frac{v_0}{\alpha(X_0(u_0))} + \sum_{i=1}^{n-1} \frac{v_i}{\alpha(\widetilde{X}_i(X_0(u_0), u_1, ..., u_i))} \leq t\},$$

and

$$A_2 = \{\underline{t} \mid t - \frac{v_0}{\alpha(X_0(u_0))} - \sum_{i=1}^{n-1} \frac{v_i}{\alpha(\widetilde{X}_i(X_0(u_0), u_1, ..., u_i))} < \frac{v_n}{\alpha(\mathcal{X}')}\}.$$

Hence

$$P(X(t) = \mathcal{X}')$$

$$= \sum_{n=0}^{\infty} \int_{(0,1]^{n+1}} \int_{(0,\infty)^n} \left[\int_0^\infty 1_{A_1}(\underline{t}) 1_{A_2}(\underline{t}) e^{-v_n} dv_n\right] e^{-\sum_{i=0}^{n-1} v_i} du_0, ..., dv_{n-1}$$

$$= \sum_{n=0}^{\infty} \int_{(0,1]^{n+1}} \int_{(0,\infty)^n} \left[\int_0^\infty 1_{A_2}(\underline{t}) e^{-v_n} dv_n\right] 1_{A_1}(\underline{t}) e^{-\sum_{i=0}^{n-1} v_i} du_0, ..., dv_{n-1}$$

$$= \sum_{n=0}^{\infty} \int_{(0,1]^{n+1}} \int_{(0,\infty)^n} \left[\int_{\alpha(\mathcal{X}')(t-T_n)}^\infty e^{-v_n} dv_n\right] 1_{A_1}(\underline{t}) e^{-\sum_{i=0}^{n-1} v_i} du_0, ..., dv_{n-1}$$

$$= \sum_{n=0}^{\infty} \int_{(0,1]^{n+1}} \int_{(0,\infty)^n} e^{-\alpha(\mathcal{X}')(t-T_n)} 1_{A_1}(\underline{t}) e^{-\sum_{i=0}^{n-1} v_i} du_0, ..., dv_{n-1}.$$

Similarly, we have

$$P(X(t+s) = \mathcal{X}, X(t) = \mathcal{X}')$$

$$= \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \int_{(0,1]^{n+l+1}} \int_{(0,\infty)^{n+l+1}} P(X_{n+l} = \mathcal{X}, T_{n+l} \leq t+s < T_{n+l+1},$$

$$X_n = \mathcal{X}', T_n \leq t < T_{n+1} \mid U_i = u_i,$$

$$V_i = v_i, 0 \leq i \leq n+l)$$

$$e^{-\sum_{i=0}^{n+l} v_i} du_0, ..., dv_{n+l}$$

$$= \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \int_{(0,1]^{n+l+1}} \int_{(0,\infty)^{n+l+1}} 1_{A'}(\underline{t}) e^{-\sum_{i=0}^{n+l} v_i} du_0, ..., dv_{n+l},$$

where

$$A' = \{\underline{t} \mid \widetilde{X}_l(\mathcal{X}', u_{n+1}, ..., u_{n+l}) = \mathcal{X}, \widetilde{X}_n(X_0(u_0), u_1, ..., u_n) = \mathcal{X}',$$

$$\widetilde{T}_l(\mathcal{X}', 0, u_{n+1}, ..., u_{n+l-1}, v_n, ..., v_{n+l-1}) +$$

$$\widetilde{T}_n(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}) \le t + s,$$

$$[t - \widetilde{T}_n(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1})] +$$

$$[s - \widetilde{T}_l(\mathcal{X}', 0, u_{n+1}, ..., u_{n+l-1}, v_n, ..., v_{n+l-1})] < \frac{v_{n+l}}{\alpha(\mathcal{X})},$$

$$\widetilde{T}_n(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}) \le t,$$

$$t - \widetilde{T}_n(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}) < \frac{v_n}{\alpha(\mathcal{X}')}\},$$

and

$$\widetilde{T}_l(\mathcal{X}', 0, u_{n+1}, ..., u_{n+l-1}, v_n, ..., v_{n+l-1}) = \frac{v_n}{\alpha(\mathcal{X}')} + \sum_{i=1}^{l-1} \frac{v_{n+i}}{\alpha(\widetilde{X}_i(\mathcal{X}', u_{n+1}, ..., u_{n+i}))}.$$

Hence

$$P(X(t+s) = \mathcal{X}, X(t) = \mathcal{X}')$$

$$= \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \int_{(0,1]^{n+l+1}} \int_{(0,\infty)^{n+l}} \left[\int_0^{\infty} 1_{A'}(\underline{t}) e^{-v_n} \, dv_n \right]$$

$$e^{-\sum_{i=0}^{n-1} v_i} e^{-\sum_{i=1}^{l} v_{n+i}}$$

$$du_0, ..., dv_{n-1}, dv_{n+1}, ..., dv_{n+l}.$$

Change variables

$$\frac{v_n''}{\alpha(\mathcal{X}')} = \frac{v_n}{\alpha(\mathcal{X}')} + \widetilde{T}_n(X_0(u_0), 0, u_1, ..., u_{n-1}, v_0, ..., v_{n-1}) - t,$$

$$= \frac{v_n}{\alpha(\mathcal{X}')} + \frac{v_0}{\alpha(X_0(u_0))} + \sum_{i=1}^{n-1} \frac{v_i}{\alpha(\widetilde{X}_i(X_0(u_0), u_1, ..., u_i))} - t,$$

then $dv_n'' = dv_n$, and

$$\int_0^{\infty} 1_{A'}(..., v_n, ...) e^{-v_n} \, dv_n$$

$$= e^{-\alpha(\mathcal{X}')(t-T_n)} \int_{\alpha(\mathcal{X}')(T_n-t)}^{\infty} 1_{A'}(..., v_n'' + \alpha(\mathcal{X}')(t - T_n), ...) e^{-v_n''} \, dv_n''. \quad (5.10)$$

65

Since A contains the condition that $\frac{v_n''}{\alpha(\mathcal{X}')} = \frac{v_n}{\alpha(\mathcal{X}')} + T_n - t > 0$, then the integrand vanishes unless $\frac{v_n''}{\alpha(\mathcal{X}')} > 0$. Hence (5.10) can be written as

$$e^{-\alpha(\mathcal{X}')(t-T_n)} \int_0^\infty 1_{A'}(..., v_n'' + \alpha(\mathcal{X}')(t - T_n), ...)e^{-v_n''}\, dv_n''.$$

Then A' can also be split into two parts: $A_1' = A_1$ and

$$A_2' = \{\underline{t}' \mid \widetilde{X}_l(\mathcal{X}', u_{n+1}, ..., u_{n+l}) = \mathcal{X},$$

$$0 \le s - \frac{v_n''}{\alpha(\mathcal{X}')} - \sum_{i=1}^{l-1} \frac{v_{n+i}}{\alpha(\widetilde{X}_i(\mathcal{X}', u_{n+1}, ..., u_{n+i}))} < \frac{v_{n+l}}{\alpha(\mathcal{X})}\},$$

where $\underline{t}' = (u_0, ..., u_{n+l}, v_0, ..., v_{n-1}, v_n'', v_{n+1}, ..., v_{n+l})$.

Let $u_{n+i} = u_i'$ and $v_{n+i} = v_i'$ for $1 \le i \le l$, and $v_n'' = v_0'$, then

$$A_2' = \{\underline{t}' \mid \widetilde{X}_l(\mathcal{X}', u_1', ..., u_l') = \mathcal{X},$$

$$0 \le s - \frac{v_0'}{\alpha(\mathcal{X}')} - \sum_{i=1}^{l-1} \frac{v_i'}{\alpha(\widetilde{X}_i(\mathcal{X}', u_1', ..., u_i'))} < \frac{v_l'}{\alpha(\mathcal{X})}\},$$

Hence

$$P(\, X(t + s) = \mathcal{X}, X(t) = \mathcal{X}')$$
$$= \Big[\sum_{n=0}^\infty \int_{(0,1]^{n+1}} \int_{(0,\infty)^n} e^{-\alpha(\mathcal{X}')(t-T_n)} 1_{A_1'}(..., v_0' + \alpha(\mathcal{X}')(t - T_n), ...)$$
$$e^{-\sum_{i=0}^{n-1} v_i}\, du_0, ..., du_n, dv_0, ..., dv_{n-1}\Big]$$
$$\Big[\sum_{l=0}^\infty \int_{(0,1]^l} \int_{(0,\infty)^l} \big[\int_0^\infty 1_{A_2'}(..., v_0' + \alpha(\mathcal{X}')(t - T_n), ...)e^{-v_0'}\, dv_0'\big]$$
$$e^{-\sum_{i=1}^{l} v_i'}\, du_1', ..., du_l', dv_1', ..., dv_l'\Big]$$
$$= P(X(t) = \mathcal{X}') \sum_{l=0}^\infty \int_{(0,1]^l} \int_{(0,\infty)^{l+1}} 1_{A_2'}(..., v_0' + \alpha(\mathcal{X}')(t - T_n), ...)$$
$$e^{-\sum_{i=0}^{l} v_i'}\, du_1', ..., du_l', dv_0', ..., dv_l'. \qquad (5.11)$$

66

We have

$$P(X(s) = \mathcal{X} \mid X(0) = \mathcal{X}') = \frac{P(X(s) = \mathcal{X}, X(0) = \mathcal{X}')}{P(X_0 = \mathcal{X}')}$$
$$= \frac{P(X(s) = \mathcal{X}, X(0) = \mathcal{X}')}{f_{\mathcal{X}'}(0)}.$$

We can obtain the similar representation using the above approach for the numerator $P(X(s) = \mathcal{X}, X(0) = \mathcal{X}')$. For this representation of $P(X(s) = \mathcal{X}, X(0) = \mathcal{X}')$, since $X(0) = \mathcal{X}'$ is given, the integrand integrating on du_0 vanishes unless u_0 is in a unique subinterval of $(0,1]$ with length $f_{\mathcal{X}'}(0)$ by our construction. Hence the integrand integrating on du_0 is equal to the constant $f_{\mathcal{X}'}(0)$, which cancels the denominator. By the similar operations demonstrated above, we can obtain

$$P(X(s) = \mathcal{X} \mid X(0) = \mathcal{X}')$$
$$= \sum_{l=0}^{\infty} \int_{(0,1]^l} \int_{(0,\infty)^{l+1}} 1_{A_2'}(..., v_0' + \alpha(\mathcal{X}')(t - T_n), ...)$$
$$e^{-\sum_{i=0}^{l} v_i'} \, du_1', ..., du_l', dv_0', ..., dv_l'.$$

By (5.11), we have

$$P(X(t+s) = \mathcal{X} \mid X(t) = \mathcal{X}') = \frac{P(X(t) = \mathcal{X}') \cdot P(X(s) = \mathcal{X} \mid X(0) = \mathcal{X}')}{P(X(t) = \mathcal{X}')}$$
$$= P(X(s) = \mathcal{X} \mid X(0) = \mathcal{X}'),$$

i.e. the process we constructed has the stationary transition probabilities.

Assume that the state space $\mathcal{V}$ is finite. We claim that the conditional probabilities derived from our construction of the Markov process satisfies the three postulates we listed about the conditional probability to derive the Master equation, and hence the time dependent probability distribution for this Markov process satisfies the Master Equation.

We starts with the first postulate: $(\mathcal{X}', \mathcal{X}) \in E$, i.e. $\forall \mathcal{X}', \mathcal{X} = \mathcal{X}' - \mathbf{n}_r \in \mathcal{V}$, as $dt \downarrow 0$,

$$P(X(t+dt) = \mathcal{X} \mid X(t) = \mathcal{X}') = W(\mathcal{X}', \mathcal{X})dt + o(dt).$$

67

By the stationary transition probability property of the constructed Markov process, we have $\forall \mathcal{X}', \mathcal{X} = \mathcal{X}' - \mathbf{n}_r \in \mathcal{V}$,

$$P(X(t + dt) = \mathcal{X} \mid X(t) = \mathcal{X}') = P(X(dt) = \mathcal{X} \mid X(0) = \mathcal{X}'). \qquad (5.12)$$

The state of the process is $\mathcal{X}$ at time dt given the state is $\mathcal{X}'$ at time 0 indicates that the holding time of state $X_0 = \mathcal{X}'$ is less than or equal to dt, and this probability is

$$P(\frac{V_0}{\alpha(X_0)} < dt \mid X_0 = \mathcal{X}') = 1 - e^{-\alpha(\mathcal{X}')dt} = \alpha(\mathcal{X}')dt + o(dt)$$

since $\frac{V_0}{\alpha(\mathcal{X}')}$ is exponentially distributed with parameter $\alpha(\mathcal{X}')$ by our construction.

Since

$$P(X(dt) = \mathcal{X} \mid X(0) = \mathcal{X}') = \sum_{n=0}^{\infty} P(X_n = \mathcal{X}, T_n \leq dt < T_{n+1} \mid X_0 = \mathcal{X}'),$$

we have to reach state $\mathcal{X} = \mathcal{X}' - \mathbf{n}_r$ from $\mathcal{X}'$ in time interval $(0, dt]$ as $dt \downarrow 0$, we have the following ways:

Through one reaction r in $(0, dt]$ where r connects the states $\mathcal{X}'$ and $\mathcal{X}$, i.e. $n = 1$. This probability is equal to the joint probability that the holding time of state $\mathcal{X}'$ is less than or equal to dt, and the only reaction occurs during that time interval is r, which is $\frac{W(\mathcal{X}', \mathcal{X})}{\alpha(\mathcal{X}')}$ by the construction. Hence this joint probability is

$$[\alpha(\mathcal{X}')dt + o(dt)] \cdot \frac{W(\mathcal{X}', \mathcal{X})}{\alpha(\mathcal{X}')} = W(\mathcal{X}', \mathcal{X})dt + o(dt);$$

Through two reactions in $(0, dt]$: the state of the process starts from $\mathcal{X}'$ to some $\mathcal{X}''$ then to $\mathcal{X}$ in $(0, dt]$, i.e. $n = 2$. For the first step, the process can reach less than or equal to $|R|$ different states from $\mathcal{X}'$ through the first reaction in the time duration less than dt, and the corresponding probability is less than $|R| \cdot [\alpha(\mathcal{X}') \cdot dt + o(dt)]$; Step 2 is that from these $|R|$ states we must reach the state $\mathcal{X}$ through the second

reaction in the time interval less than dt, and the corresponding probability is less than $\alpha \cdot dt + o(dt)$, where $\alpha = \max_{\mathcal{X} \in \mathcal{V}} \alpha(\mathcal{X})$. Hence the probability that the process reaches state $\mathcal{X}$ from $\mathcal{X}'$ through two reactions in time interval $(0, dt]$ is less than:

$$|R| \cdot [\alpha \cdot dt + o(dt)]^2;$$

Through three reactions in $(0, dt]$, , i.e. $n = 3$. Using the same idea as above, we obtain this probability is less than:

$$|R|^2 \cdot [\alpha \cdot dt + o(dt)]^3;$$

Continue in this way we can obtain that the probability to reach state $\mathcal{X}$ from $\mathcal{X}'$ in time interval $(0, dt]$ through $k + 1$ reactions is less than:

$$|R|^k \cdot [\alpha \cdot dt + o(dt)]^{k+1},$$

where $k \geq 1$.

As $dt \downarrow 0$,

$$\sum_{k=1}^{\infty} |R|^k [\alpha \cdot dt + o(dt)]^{k+1} = \alpha \cdot dt \sum_{k=1}^{\infty} [|R| \cdot \alpha \cdot dt]^k + o(dt)$$

$$= \frac{|R| \cdot \alpha^2 \cdot dt^2}{1 - |R| \cdot \alpha \cdot dt} + o(dt)$$

$$= o(dt).$$

Hence we obtain the probability to reach state $\mathcal{X}$ from $\mathcal{X}'$ in time interval $(0, dt]$ is

$$[W(\mathcal{X}', \mathcal{X})dt + o(dt)] + o(dt) = W(\mathcal{X}', \mathcal{X})dt + o(dt),$$

i.e., $\forall \mathcal{X}', \mathcal{X} = \mathcal{X}' - \mathbf{n}_r \in \mathcal{V}$, as $dt \downarrow 0$,

$$P(X(t + dt) = \mathcal{X} \mid X(t) = \mathcal{X}') = W(\mathcal{X}', \mathcal{X})dt + o(dt).$$

To prove the second postulate

$$P(X(t + dt) = \mathcal{X} \mid X(t) = \mathcal{X}) = 1 - \alpha(\mathcal{X})dt + o(dt),$$

it suffices to prove that

$$P(X(dt) = \mathcal{X} \mid X(0) = \mathcal{X}) = 1 - \alpha(\mathcal{X})dt + o(dt).$$

We observe that there are two possible ways to be at state $\mathcal{X}$ at time dt given $X(0) = \mathcal{X}$: no reaction occurs in the time interval $(0, dt]$, whose probability is

$$P(\frac{V_0}{\alpha(X_0)} > dt \mid X_0 = \mathcal{X}) = e^{-\alpha(\mathcal{X})dt} + o(dt) = 1 - \alpha(\mathcal{X})dt + o(dt);$$

or more than one reaction occur in $(0, dt]$ to reach $\mathcal{X}$ from $\mathcal{X}$, and this probability is $o(dt)$ by the proof of postulate 1.

Hence, $\forall \mathcal{X} \in \mathcal{V}$, as $dt \downarrow 0$,

$$P(X(t + dt) = \mathcal{X} \mid X(t) = \mathcal{X}) = 1 - \alpha(\mathcal{X})dt + o(dt).$$

Last, we verify the third postulate. $\forall \mathcal{X}, \mathcal{X}' \in \mathcal{V}$, where $(\mathcal{X}', \mathcal{X}) \notin E \cup \Delta$, the process reaches state $\mathcal{X}$ at time dt from state $\mathcal{X}'$ at time 0 implies that the holding time of state $X_0 = \mathcal{X}'$ is less than dt. Hence $\forall \mathcal{X} \in \mathcal{V}$, for some $\epsilon > 0$,

$$\begin{aligned}
&\sup\{\frac{P(X(t + dt) = \mathcal{X} \mid X(t) = \mathcal{X}')}{dt} \mid 0 < dt < \epsilon, (\mathcal{X}', \mathcal{X}) \notin E \cup \Delta\} \\
&= \sup\{\frac{P(X(dt) = \mathcal{X} \mid X(0) = \mathcal{X}')}{dt} \mid 0 < dt < \epsilon, (\mathcal{X}', \mathcal{X}) \notin E \cup \Delta\} \\
&< \sup\{\frac{P(\frac{V_0}{\alpha(X_0)} < dt \mid X_0 = \mathcal{X}')}{dt} \mid 0 < dt < \epsilon, (\mathcal{X}', \mathcal{X}) \notin E \cup \Delta\} \\
&= \sup\{\frac{\alpha(\mathcal{X}')dt + o(dt)}{dt} \mid 0 < dt < \epsilon, (\mathcal{X}', \mathcal{X}) \notin E \cup \Delta\} \\
&< \sup\{\alpha(\mathcal{X}') \mid (\mathcal{X}', \mathcal{X}) \notin E \cup \Delta\} + \epsilon \\
&< \infty
\end{aligned}$$

since the state space $\mathcal{V}$ is finite, i.e. $\alpha(\mathcal{X}') < \infty \ \forall \mathcal{X}' \in \mathcal{V}$.

Chapter 6

Derivation of the Model

6.1 The Auto-Regulating Genetic Model

We will use the following notation in this section and later:

If A and B are two macromolecules, then $A \cdot B$ denotes a complex comprised of one A and one B molecule bound together by hydrogen bounds, other non-covalent electrostatic attractions and hydrophobic effects; $A \cdot B \cdot DNA_{(m,n)}$ denotes a complex of A, B and DNA, where A is bound at site m and B at site n on the DNA strand.

Figure 6.1 is a representative auto-regulating prokaryotic genetic circuit, where the protein P is encoded by gene p, and the protein dimer $P \cdot P = P_2$ auto-controls the expression of p.
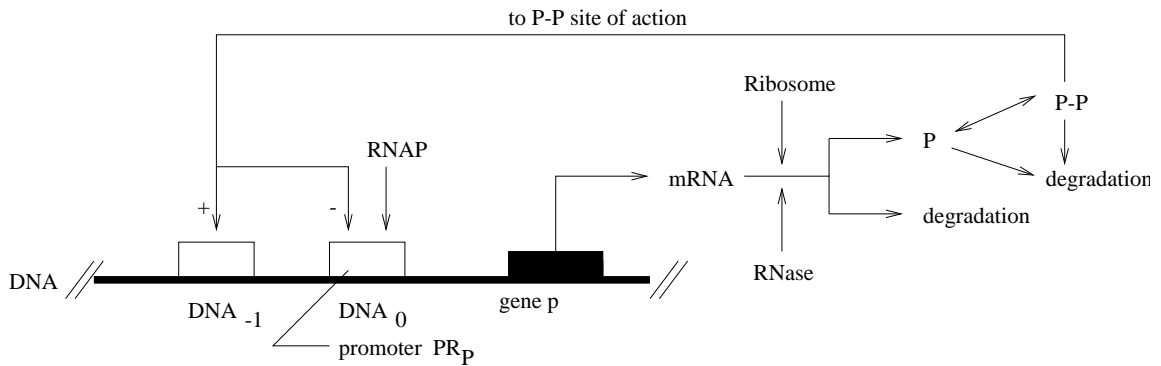

Figure 6.1: Mechanism of Auto-Regulating Prokaryotic Genetic Circuit

The mechanism for information transmission in the autoregulatory genetic network is as follows: $RNAP$ combines with the promoter PR_P at position DNA_0 in a closed form $RNAP_c \cdot DNA_0$, where the based pairs on DNA are not separated. After $RNAP$ separates the paired bases, i.e., the closed complex becomes the open complex $RNAP_o \cdot DNA_0$, transcription of gene p can be initiated. Later, the transcription processes produce a transcript $mRNA$ of gene p, which can either be translated to form protein P when combined with Ribosome, or can be degraded when acted upon by RNase. Protein P accumulates if its production rate exceeds its degradation rate. The dimerization process from P to P_2 is reversible.

The active form of a regulating protein is commonly a multimer, which in our case is the dimer P_2. P_2 regulates the expression of gene p in the following way: P_2 has a much higher affinity for the binding site DNA_{-1} than for the site DNA_0. $RNAP$ only binds at DNA_0 to yield $RNAP_c \cdot DNA_0$. However the isomerization process from $RNAP_c \cdot DNA_0$ to $RNAP_o \cdot DNA_0$ is slow. On the other hand, when P_2 binds to DNA_{-1}, it can facilitate the closed complex $P_2 \cdot RNAP_c \cdot DNA_{-1,0}$ isomerizing into the open form $P_2 \cdot RNAP_o \cdot DNA_{-1,0}$, thereby increasing the rate of transcription of gene p. This positive auto-regulating effect then leads to the accumulation of P and of P_2. As the concentration of P_2 increases, the chance that P_2 binds to the promoter at DNA_0 increases, which prevents $RNAP$ from gaining access to the promoter and hence blocks the initiation of transcription of gene p.

In addition we assume the system has the following properties:

(1) Each binding site binds one protein at a time.

(2) $RNAP$ and $Ribosome$ can only bind to DNA and $mRNA$ respectively when there is no other $RNAP$ and $Ribosome$ already bound to them.

(3) P_2 can bind at both the binding sites DNA_{-1} and DNA_0, while $RNAP$ only binds at DNA_0.

(4) P_2 can bind to DNA_{-1} even when there is already an $RNAP$ on DNA_0; simi-larly, $RNAP$ can bind to DNA_0 even when there is already a P_2 on DNA_{-1}.

More discussion on the gene regulation system can be found in ([19]).

6.2   Petri Net and Master Equation for the Model

The processes of gene expression consists of coupled chemical reactions. We consider the cell as a spatially homogeneous chemical system, which has a fixed volume con-taining many chemical species which can interreact through some specified chemical reaction channels.

Suppose the length of gene p is M nucleotides and the length of $mRNA$ is N codons. The auto-regulating genetic circuit involves the following chemical reaction species:

Chemical Species:

s_1: DNA

s_2: $RNAP$

s_3: $mRNA$

s_4: $Ribosome$

s_5: $RNase$

s_6: P

s_7: P_2

s_8: $P_2 \cdot DNA_{-1}$

s_9: $P_2 \cdot DNA_0$

s_{10}: $RNAP_c \cdot DNA_0$

s_{11}: $P_2 \cdot RNAP_c \cdot DNA_{(-1,0)}$

s_{12+m}: $\qquad RNAP_o \cdot DNA_m \quad (m = 0, ..., M)$

s_{13+M+m}: $\qquad P_2 \cdot RNAP_o \cdot DNA_{(-1,m)} \quad (m = 0, ..., M)$

$s_{14+2M+n}$: $\qquad Ribosome \cdot mRNA_n \quad (n = 0, ..., N)$.

Suppose the system would reach state $\mathcal{X}$ at a particular time, then the parameters $W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r)$ and $W(\mathcal{X} + \mathbf{n}_r, \mathcal{X})$ would have specified values at state $\mathcal{X}$ for each reaction r involved in the system. In the following we will list the chemical reactions and the corresponding values for $W(\mathcal{X}, \mathcal{X} - \mathbf{n}_r)$ and $W(\mathcal{X} + \mathbf{n}_r, \mathcal{X})$ using (3.2) (for simplicity, only the states of reactants and products are listed for each reaction):

Transcription Reactions:

Initiation:

r_1 :

$DNA + RNAP \longrightarrow RNAP_c \cdot DNA_0$

$s_1 \quad + s_2 \quad \longrightarrow s_{10}$
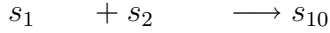
$\mathcal{X} + \mathbf{n}_{r_1} = (\mathcal{X}(s_1) + 1, \mathcal{X}(s_2) + 1, \mathcal{X}(s_{10}) - 1)$

$l_{r_1} = 2$

$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_1}) = C(r_1)\mathcal{X}(s_1)\mathcal{X}(s_2)/V$

$W(\mathcal{X} + \mathbf{n}_{r_1}, \mathcal{X}) = C(r_1)(\mathcal{X}(s_1) + 1)(\mathcal{X}(s_2) + 1)/V$
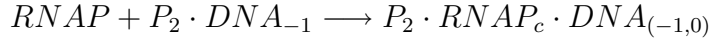
r_2 :

$RNAP_c \cdot DNA_0 \longrightarrow RNAP_o \cdot DNA_0$

$s_{10} \qquad\qquad \longrightarrow s_{12}$

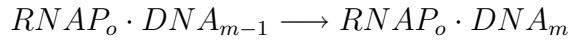$\mathcal{X} + \mathbf{n}_{r_2} = (\mathcal{X}(s_{10}) + 1, \mathcal{X}(s_{12}) - 1)$

$l_{r_2} = 1$

$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_2}) = C(r_2)\mathcal{X}(s_{10})$

$W(\mathcal{X} + \mathbf{n}_{r_2}, \mathcal{X}) = C(r_2)(\mathcal{X}(s_{10}) + 1)$

r_3 :

$RNAP + P_2 \cdot DNA_{-1} \longrightarrow P_2 \cdot RNAP_c \cdot DNA_{(-1,0)}$

$s_2 \qquad + s_8 \qquad\qquad \longrightarrow s_{11}$

$\mathcal{X} + \mathbf{n}_{r_3} = (\mathcal{X}(s_2) + 1, \mathcal{X}(s_8) + 1, \mathcal{X}(s_{11}) - 1)$

$l_{r_3} = 2$

$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_3}) = C(r_3)\mathcal{X}(s_2)\mathcal{X}(s_8)/V$

$W(\mathcal{X} + \mathbf{n}_{r_3}, \mathcal{X}) = C(r_3)(\mathcal{X}(s_2) + 1)(\mathcal{X}(s_8) + 1)/V$

r_4 :

$P_2 \cdot RNAP_c \cdot DNA_{(-1,0)} \longrightarrow P_2 \cdot RNAP_o \cdot DNA_{(-1,0)}$

$s_{11} \qquad\qquad\qquad\qquad \longrightarrow s_{13+M}$
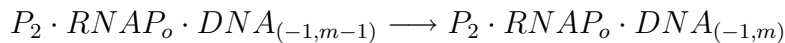
$\mathcal{X} + \mathbf{n}_{r_4} = (\mathcal{X}(s_{11}) + 1, \mathcal{X}(s_{13+M}) - 1)$

$l_{r_4} = 1$

$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_4}) = C(r_4)\mathcal{X}(s_{11})$

$W(\mathcal{X} + \mathbf{n}_{r_4}, \mathcal{X}) = C(r_4)(\mathcal{X}(s_{11}) + 1)$

Elongation: $\qquad (m = 1, ..., M)$

r_{4+m} :

$RNAP_o \cdot DNA_{m-1} \longrightarrow RNAP_o \cdot DNA_m$

$s_{11+m} \qquad\qquad \longrightarrow s_{12+m}$

$\mathcal{X} + \mathbf{n}_{r_{4+m}} = (\mathcal{X}(s_{11+m}) + 1, \mathcal{X}(s_{12+m}) - 1)$

$l_{r_{4+m}} = 1$

$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{4+m}}) = C(r_{4+m})\mathcal{X}(s_{11+m})$

$W(\mathcal{X} + \mathbf{n}_{r_{4+m}}, \mathcal{X}) = C(r_{4+m})(\mathcal{X}(s_{11+m}) + 1)$

r_{4+M+m} :

$P_2 \cdot RNAP_o \cdot DNA_{(-1,m-1)} \longrightarrow P_2 \cdot RNAP_o \cdot DNA_{(-1,m)}$

$$s_{12+M+m} \longrightarrow s_{13+M+m}$$

$$\mathcal{X} + \mathbf{n}_{r_{4+M+m}} = (\mathcal{X}(s_{12+M+m}) + 1, \mathcal{X}(s_{13+M+m}) - 1)$$

$$l_{r_{4+M+m}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{4+M+m}}) = C(r_{4+M+m})\mathcal{X}(s_{12+M+m})$$

$$W(\mathcal{X} + \mathbf{n}_{r_{4+M+m}}, \mathcal{X}) = C(r_{4+M+m})(\mathcal{X}(s_{12+M+m}) + 1)$$

Termination:

$$r_{5+2M} :$$

$$RNAP_o \cdot DNA_M \longrightarrow DNA + RNAP + mRNA$$

$$s_{12+M} \longrightarrow s_1 \quad + s_2 \quad + s_3$$

$$\mathcal{X} + \mathbf{n}_{r_{5+2M}} = (\mathcal{X}(s_1) - 1, \mathcal{X}(s_2) - 1, \mathcal{X}(s_3) - 1, \mathcal{X}(s_{12+M}) + 1)$$

$$l_{r_{5+2M}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{5+2M}}) = C(r_{5+2M})\mathcal{X}(s_{12+M})$$

$$W(\mathcal{X} + \mathbf{n}_{r_{5+2M}}, \mathcal{X}) = C(r_{5+2M})(\mathcal{X}(s_{12+M}) + 1)$$

$$r_{6+2M} :$$

$$P_2 \cdot RNAP_o \cdot DNA_{(-1,M)} \longrightarrow RNAP + mRNA + P_2 \cdot DNA_{-1}$$

$$s_{13+2M} \longrightarrow s_2 \quad + s_3 \quad + s_8$$

$$\mathcal{X} + \mathbf{n}_{r_{6+2M}} = (\mathcal{X}(s_2) - 1, \mathcal{X}(s_3) - 1, \mathcal{X}(s_8) - 1, \mathcal{X}(s_{13+2M}) + 1)$$

$$l_{r_{6+2M}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{6+2M}}) = C(r_{6+2M})\mathcal{X}(s_{13+2M})$$

$$W(\mathcal{X} + \mathbf{n}_{r_{6+2M}}, \mathcal{X}) = C(r_{6+2M})(\mathcal{X}(s_{13+2M}) + 1)$$

Translation Reactions:

Initiation:

$$r_{7+2M} :$$

$$mRNA + Ribosome \longrightarrow Ribosome \cdot mRNA_0$$

$$s_3 \quad + s_4 \quad \longrightarrow s_{14+2M}$$

$$\mathcal{X} + \mathbf{n}_{r_{7+2M}} = (\mathcal{X}(s_3) + 1, \mathcal{X}(s_4) + 1, \mathcal{X}(s_{14+2M}) - 1)$$
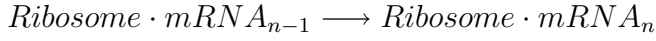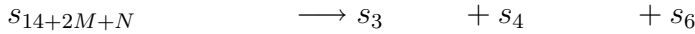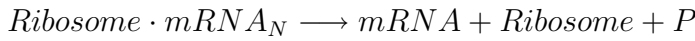
$$l_{r_{7+2M}} = 2$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{7+2M}}) = C(r_{7+2M})\mathcal{X}(s_3)\mathcal{X}(s_4)/V$$

$$W(\mathcal{X} + \mathbf{n}_{r_{7+2M}}, \mathcal{X}) = C(r_{7+2M})(\mathcal{X}(s_3) + 1)(\mathcal{X}(s_4) + 1)/V$$

Elongation: $\qquad (n = 1, ..., N)$

r_{7+2M+n} :

$$Ribosome \cdot mRNA_{n-1} \longrightarrow Ribosome \cdot mRNA_n$$

$$s_{13+2M+n} \qquad\qquad \longrightarrow s_{14+2M+n}$$

$$\mathcal{X} + \mathbf{n}_{r_{7+2M+n}} = (\mathcal{X}(s_{13+2M+n}) + 1, \mathcal{X}(s_{14+2M+n}) - 1)$$

$$l_{r_{7+2M+n}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{7+2M+n}}) = C(r_{7+2M+n})\mathcal{X}(s_{13+2M+n})$$

$$W(\mathcal{X} + \mathbf{n}_{r_{7+2M+n}}, \mathcal{X}) = C(r_{7+2M+n})(\mathcal{X}(s_{13+2M+n}) + 1)$$

Termination:

r_{8+2M+N} :

$$Ribosome \cdot mRNA_N \longrightarrow mRNA + Ribosome + P$$

$$s_{14+2M+N} \qquad\qquad \longrightarrow s_3 \quad + s_4 \quad + s_6$$

$$\mathcal{X} + \mathbf{n}_{r_{8+2M+N}} = (\mathcal{X}(s_3) - 1, \mathcal{X}(s_4) - 1, \mathcal{X}(s_6) - 1, \mathcal{X}(s_{14+2M+N}) + 1)$$

$$l_{r_{8+2M+n}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{8+2M+N}}) = C(r_{8+2M+N})\mathcal{X}(s_{14+2M+N})$$

$$W(\mathcal{X} + \mathbf{n}_{r_{8+2M+N}}, \mathcal{X}) = C(r_{8+2M+N})(\mathcal{X}(s_{14+2M+N}) + 1)$$

Other reactions:

Dimerization:

r_{9+2M+N} :

$$2P \longrightarrow P_2$$

$$2s_6 \longrightarrow s_7$$

$$\mathcal{X} + \mathbf{n}_{r_{9+2M+N}} = (\mathcal{X}(s_6) + 2, \mathcal{X}(s_7) - 1)$$

$$l_{r_{9+2M+n}} = 2$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{9+2M+N}}) = C(r_{9+2M+N})\mathcal{X}(s_6)(\mathcal{X}_6 - 1)/2V$$

$$W(\mathcal{X} + \mathbf{n}_{r_{9+2M+N}}, \mathcal{X}) = C(r_{9+2M+N})(\mathcal{X}_6 + 2)(\mathcal{X}(s_6) + 1)/2V$$

$r_{10+2M+N}$:

$$P_2 \longrightarrow 2P$$

$$s_7 \longrightarrow 2s_6$$

$$\mathcal{X} + \mathbf{n}_{r_{10+2M+N}} = (\mathcal{X}(s_6) - 2, \mathcal{X}(s_7) + 1)$$

$$l_{r_{10+2M+N}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{10+2M+N}}) = C(r_{10+2M+N})\mathcal{X}(s_7)$$

$$W(\mathcal{X} + \mathbf{n}_{r_{10+2M+N}}, \mathcal{X}) = C(r_{10+2M+N})(\mathcal{X}(s_7) + 1)$$

Degradation:

$r_{11+2M+N}$:

$$mRNA + RNase \longrightarrow RNase$$

$$s_3 \quad + s_5 \quad \longrightarrow s_5$$

$$\mathcal{X} + \mathbf{n}_{r_{11+2M+N}} = (\mathcal{X}(s_3) + 1, \mathcal{X}(s_5))$$

$$l_{r_{11+2M+N}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{11+2M+N}}) = C(r_{11+2M+N})\mathcal{X}(s_3)\mathcal{X}(s_5)$$

$$W(\mathcal{X} + \mathbf{n}_{r_{11+2M+N}}, \mathcal{X}) = C(r_{11+2M+N})(\mathcal{X}(s_3) + 1)\mathcal{X}(s_5)$$

$r_{12+2M+N}$:

$$P \longrightarrow \emptyset$$

$$s_6 \longrightarrow \emptyset$$

$$\mathcal{X} + \mathbf{n}_{r_{12+2M+N}} = (\mathcal{X}(s_6) + 1)$$

$$l_{r_{12+2M+N}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{12+2M+N}}) = C(r_{12+2M+N})\mathcal{X}(s_6)$$

$$W(\mathcal{X} + \mathbf{n}_{r_{12+2M+N}}, \mathcal{X}) = C(r_{12+2M+N})(\mathcal{X}(s_6) + 1)$$

$r_{13+2M+N}$:

$$P_2 \longrightarrow \emptyset$$

$$s_7 \longrightarrow \emptyset$$

$$\mathcal{X} + \mathbf{n}_{r_{13+2M+N}} = (\mathcal{X}(s_7) + 1)$$

$$l_{r_{13+2M+N}} = 2$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{13+2M+N}}) = C(r_{13+2M+N})\mathcal{X}(s_7)$$

$$W(\mathcal{X} + \mathbf{n}_{r_{13+2M+N}}, \mathcal{X}) = C(r_{13+2M+N})(\mathcal{X}(s_7) + 1)$$

Binding and Releasing: $(m = 1, ..., M)$

$r_{14+2M+N}$:

$$DNA + P_2 \longrightarrow P_2 \cdot DNA_{-1}$$
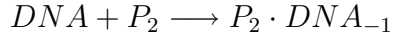
$$s_1 \quad + s_7 \longrightarrow s_8$$

$$\mathcal{X} + \mathbf{n}_{r_{14+2M+N}} = (\mathcal{X}(s_1) + 1, \mathcal{X}(s_7) + 1, \mathcal{X}(s_8) - 1)$$

$$l_{r_{14+2M+N}} = 2$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{14+2M+N}}) = C(r_{14+2M+N})\mathcal{X}(s_1)\mathcal{X}(s_7)/V$$

$$W(\mathcal{X} + \mathbf{n}_{r_{14+2M+N}}, \mathcal{X}) = C(r_{14+2M+N})(\mathcal{X}(s_1) + 1)(\mathcal{X}(s_7) + 1)/V$$

$r_{15+2M+N}$:

$$DNA + P_2 \longrightarrow P_2 \cdot DNA_0$$

$$s_1 \quad + s_7 \longrightarrow s_9$$

$$\mathcal{X} + \mathbf{n}_{r_{15+2M+N}} = (\mathcal{X}(s_1) + 1, \mathcal{X}(s_7) + 1, \mathcal{X}(s_9) - 1)$$

$$l_{r_{15+2M+N}} = 2$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{15+2M+N}}) = C(r_{15+2M+N})\mathcal{X}(s_1)\mathcal{X}(s_7)/V$$

$$W(\mathcal{X} + \mathbf{n}_{r_{15+2M+N}}, \mathcal{X}) = C(r_{15+2M+N})(\mathcal{X}(s_1) + 1)(\mathcal{X}(s_7) + 1)/V$$

$r_{16+2M+N}$:

$$P_2 \cdot DNA_{-1} \longrightarrow DNA + P_2$$

$$s_8 \qquad \longrightarrow s_1 \quad + s_7$$

$$\mathcal{X} + \mathbf{n}_{r_{16+2M+N}} = (\mathcal{X}(s_1) - 1, \mathcal{X}(s_7) - 1, \mathcal{X}(s_8) + 1)$$

$$l_{r_{16+2M+N}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{16+2M+N}}) = C(r_{16+2M+N})\mathcal{X}(s_8)$$

$$W(\mathcal{X} + \mathbf{n}_{r_{16+2M+N}}, \mathcal{X}) = C(r_{16+2M+N})(\mathcal{X}(s_8) + 1)$$

$r_{17+2M+N}$:

$$P_2 \cdot DNA_0 \longrightarrow DNA + P_2$$

$$s_9 \qquad \longrightarrow s_1 \quad + s_7$$

$$\mathcal{X} + \mathbf{n}_{r_{17+2M+N}} = (\mathcal{X}(s_1) - 1, \mathcal{X}(s_7) - 1, \mathcal{X}(s_9) + 1)$$

$$l_{r_{17+2M+N}} = 1$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{17+2M+N}}) = C(r_{17+2M+N})\mathcal{X}(s_9)$$

$$W(\mathcal{X} + \mathbf{n}_{r_{17+2M+N}}, \mathcal{X}) = C(r_{17+2M+N})(\mathcal{X}(s_9) + 1)$$

$r_{18+2M+N}$:

$$P_2 + RNAP_c \cdot DNA_0 \longrightarrow P_2 \cdot RNAP_c \cdot DNA_{(-1,0)}$$

$$s_7 + s_{10} \qquad \longrightarrow s_{11}$$

$$\mathcal{X} + \mathbf{n}_{r_{18+2M+N}} = (\mathcal{X}(s_7) + 1, \mathcal{X}(s_{10}) + 1, \mathcal{X}(s_{11}) - 1)$$

$$l_{r_{118+2M+N}} = 2$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{18+2M+N}}) = C(r_{18+2M+N})\mathcal{X}(s_7)\mathcal{X}(s_{10})/V$$

$$W(\mathcal{X} + \mathbf{n}_{r_{18+2M+N}}, \mathcal{X}) = C(r_{18+2M+N})(\mathcal{X}(s_7) + 1)(\mathcal{X}(s_{10}) + 1)/V$$
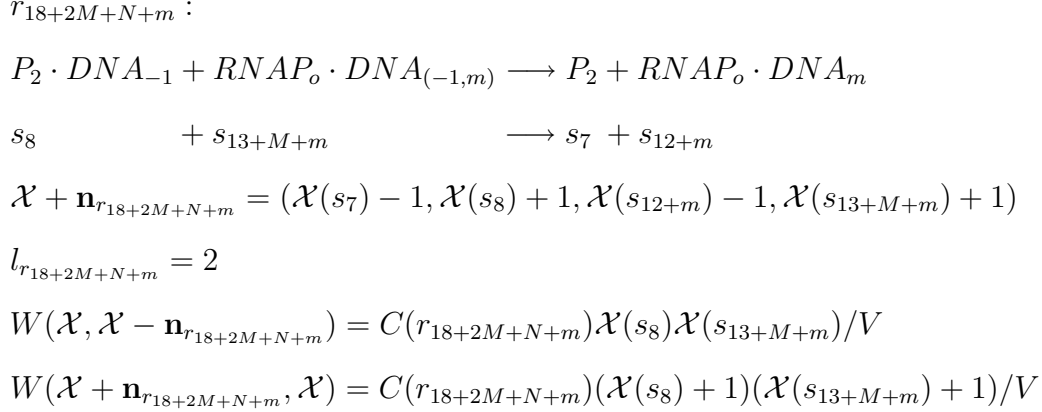
$r_{18+2M+N+m}$:

$$P_2 \cdot DNA_{-1} + RNAP_o \cdot DNA_{(-1,m)} \longrightarrow P_2 + RNAP_o \cdot DNA_m$$

$$s_8 \qquad + s_{13+M+m} \qquad \longrightarrow s_7 + s_{12+m}$$

$$\mathcal{X} + \mathbf{n}_{r_{18+2M+N+m}} = (\mathcal{X}(s_7) - 1, \mathcal{X}(s_8) + 1, \mathcal{X}(s_{12+m}) - 1, \mathcal{X}(s_{13+M+m}) + 1)$$

$$l_{r_{18+2M+N+m}} = 2$$

$$W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_{18+2M+N+m}}) = C(r_{18+2M+N+m})\mathcal{X}(s_8)\mathcal{X}(s_{13+M+m})/V$$

$$W(\mathcal{X} + \mathbf{n}_{r_{18+2M+N+m}}, \mathcal{X}) = C(r_{18+2M+N+m})(\mathcal{X}(s_8) + 1)(\mathcal{X}(s_{13+M+m}) + 1)/V$$

Fig 6.2 gives the Petri net for this system in the case when $M = 3$, $N = 2$.

Now we can apply the Master Equation approach to the model using the values listed above for $W(\mathcal{X} + \mathbf{n}_{r_j}, \mathcal{X})$ and $W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_j})$ for $1 \leq j \leq 18 + 3M + N$:

$$\frac{d}{dt} f_\mathcal{X}(t) = \sum_{j=1}^{18+3M+N} W(\mathcal{X} + \mathbf{n}_{r_j}, \mathcal{X}) f_{\mathcal{X}+\mathbf{n}_{r_j}}(t) - [\sum_{j=1}^{18+3M+N} W(\mathcal{X}, \mathcal{X} - \mathbf{n}_{r_j})] f_\mathcal{X}(t) \quad (6.1)$$

where $f_\mathcal{X}(t)$ is the probability that the state of the system is $\mathcal{X}$ at time t.

([1]) and ([16]) discussed how to use the stochastic model to simulate the developmental pathway bifurcation in a more complicated gene regulating system: a Phage λ-infected Escherichia Coli cell.
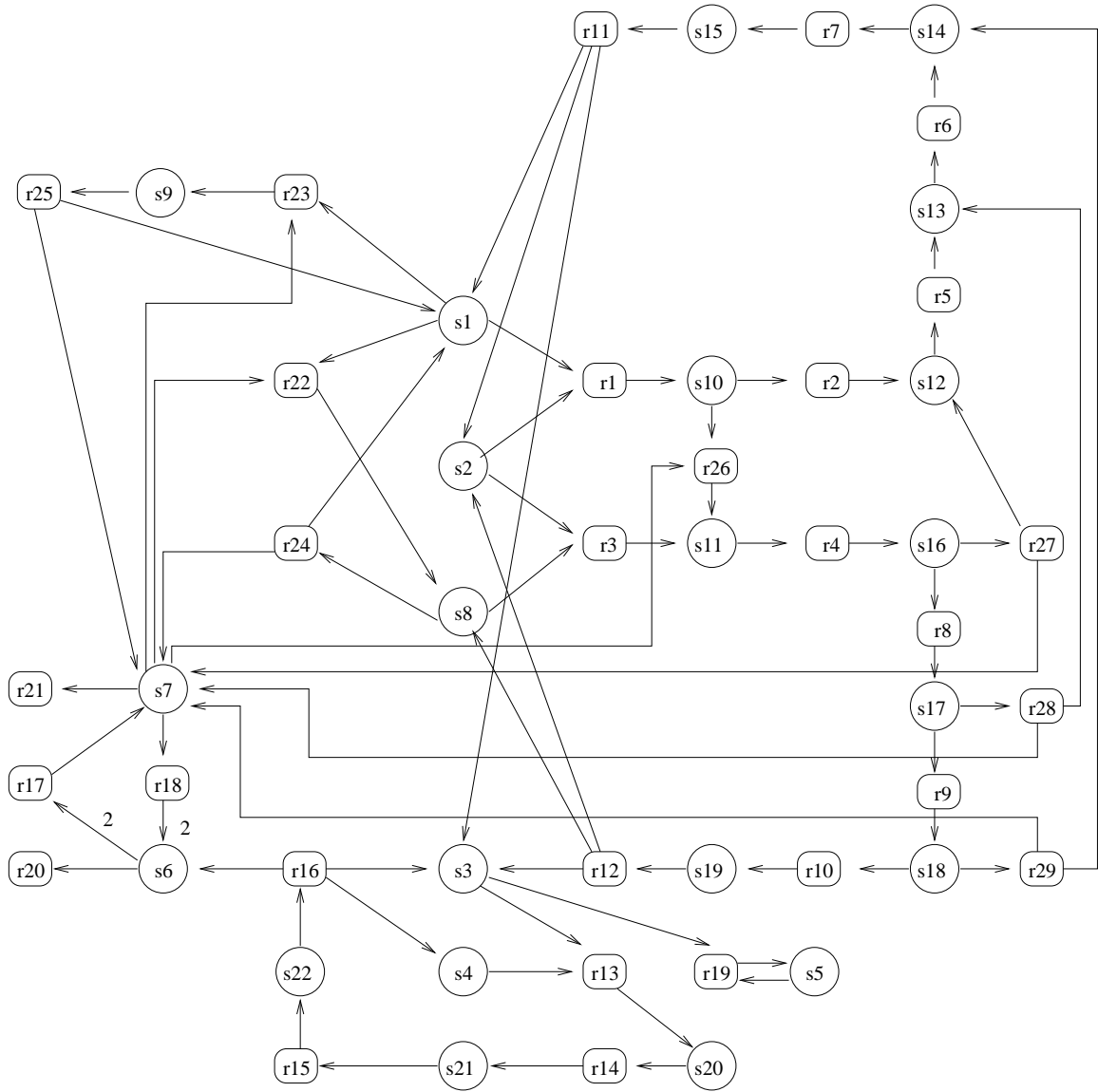
Figure 6.2: Petri Net for the Model when $M = 3$, $N = 2$.

Chapter 7

Simulation of Models Using the Gillespie Algorithm

7.1 Introduction of the Simulation Software and Gillespie Algorithm

In this chapter, we will use a biochemical simulation software called Jarnac (version 1.19) ([22]) and the Gillespie algorithm to simulate the stochastic time evolution of a simplified model of the auto-regulating genetic circuit.

Jarnac is a language for describing and manipulating cellular system models and can be used to describe any physical system which can be described in terms of a network and associated flows including gene networks. It can support deterministic as well as stochastic simulation of a biochemical network, perform time course or steady state based simulations.

To input the model which we want to simulate, we have to first define the chemical system in Jarnac. For a gene network, the system is usually a cell. Each chemical reaction in Jarnac is specified by the species involved in the reaction, the stoichiometric details of the reaction and the reaction rate of the reaction. The following is an example to input a gene network involving the reactions $2s_1 \rightarrow s_2$ and $s_2 \rightarrow 2s_1$

in Jarnac, where the reaction rate for the first reaction is 1 and for the second is 0.5:

$$p = \text{defn cell}$$

$$2s_1 \rightarrow s_2; k1;$$

$$s_2 \rightarrow 2s_1; k2;$$

$$\text{end};$$

$$p.k1 = 1;$$

$$p.k2 = 0.5;$$

The stochastic simulation algorithm we will use in Jarnac is called the **Gillespie algorithm**. The Gillespie algorithm is a Monte Carlo simulation technique, which use the same idea as the construction of the Markov process. The Gillespie algorithm and the Master equation approach are equivalent in the sense that both are consequences of the fundamental hypothesis (3.1) ([9]). ([5]) gives a modified Gillespie algorithm to simulate the genetic networks. The following introduction of the algorithm is given by ([8]) and ([9]).

To simulate the time evolution of a chemical system with P species and Q reactions by the Gillespie algorithm, the steps are:

Step 0. Input the values of the reaction constants for the Q reactions and the initial state of the system, i.e., the initial number of molecules of P species. Set the time variable $t = 0$. Specify the stopping time t_{stop}.

Step 1. Calculate and store the number of distinct molecular combinations for each reaction at current state.

Step 2. Generate two random numbers using a unit-interval uniform random number generator, and calculate which reaction will occur next and the holding time of the current state as we construct the Markov process.

Step 3. Increase t by the holding time of current state, and update the numbers of molecules of the reactants and products involved in the reaction which will occur next, to reflect the occurrence of this reaction.

Step 4. If $t > t_{stop}$, or if no more reactants remain, terminate the simulation; otherwise, return to step 1.

To perform the Gillespie algorithm of the model p defined earlier in Jarnac, we use the following script:

$$m = \text{gillespie}(p, 1000, [< p.\text{Time} >, < p.s_1 >, < p.s_2 >]);$$

$$\text{graph}(m);$$

where the simulation runs from time 0 to 1000 seconds, and the command $\text{graph}(m)$ will give us the outcome of time evolution of numbers of molecules of s_1 and s_2 obtained from the simulation.

7.2 Petri Net of a Simplified Model

In the simplified auto-regulating genetic system, we combine the transcription and translation processes as a single protein synthesis process. The model contains a single copy of gene, which is initially inactive, but which may be activated subsequently by a initiation process. The activated gene can produce protein which may be degraded at any time. The dimerization reaction of protein is reversible. The dimer protein can either facilitate or block the initiation process by binding to different position of the DNA strand.

We use the following chemical species and reactions to represent the simplified model:

Chemical Species:

s_1 : inactive gene

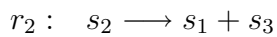s_2 : active gene

s_3 : protein

s_4 : protein dimer
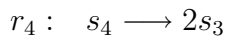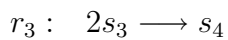
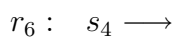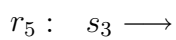s_5 : blocked gene

Chemical Reactions:

Slow Initiation:

r_1 : $s_1 \longrightarrow s_2$

Gene Expression:

r_2 : $s_2 \longrightarrow s_1 + s_3$

Dimerization:

r_3 : $2s_3 \longrightarrow s_4$

r_4 : $s_4 \longrightarrow 2s_3$

Degradation:

r_5 : $s_3 \longrightarrow$

r_6 : $s_4 \longrightarrow$

Fast Initiation:

r_7 : $s_1 + s_4 \longrightarrow s_2 + s_4$
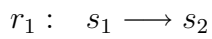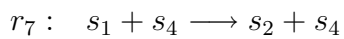
Blocking Process:

r_8 : $s_1 + s_4 \longrightarrow s_5$
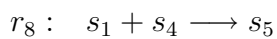
Releasing Process:

r_9 : $s_5 \longrightarrow s_1 + s_4$

The Petri net representation of the simplified model of the auto-regulating genetic circuit is shown in Fig (7.1).
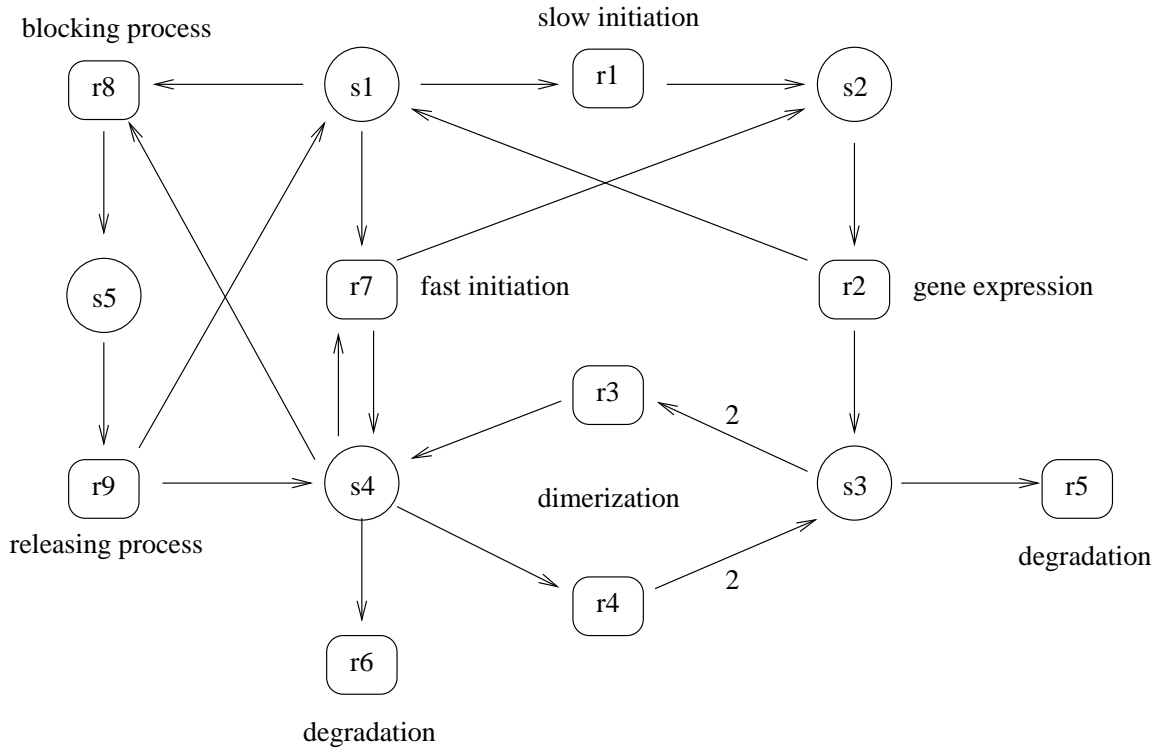
Figure 7.1: Petri Net for the Simplified Model.

7.3 Simulation Outcomes

Let the rates for reactions from r_1 to r_9 be 0.1, 0.1, 0.04, 0.02, 0.002, 0.002, 0.3, 0.02 and 0.05, which are adapted from ([1]). We will use the following script to simulate the model at an initial state $(s_1, s_2, s_3, s_4, s_5) = (1, 0, 0, 0, 0)$ from time 0 to 10000 seconds:

$p = $ defn cell

$s_1 \rightarrow s_2; k1;$

$s_2 \rightarrow s_1 + s_3; k2;$

$2s_3 \rightarrow s_4; k3;$

$s_4 \rightarrow 2s_3; k4;$

$s_3 \rightarrow; k5;$

$s_4 \rightarrow; k6;$

$s_1 + s_4 \rightarrow s_2 + s_4; k7;$

$s_1 + s_4 \rightarrow s_5; k8;$

$s_5 \rightarrow s_1 + s_4; k9;$

end;

$p.k_1 = 0.1;$

$p.k_2 = 0.1;$

$p.k_3 = 0.04;$

$p.k_4 = 0.02;$

$p.k_5 = 0.002;$

$p.k_6 = 0.002;$

$p.k_7 = 0.3;$

$p.k_8 = 0.02;$

$p.k_9 = 0.05;$

$p.s_1 = 1;$

$p.s_2 = 0;$

$p.s_3 = 0;$

$p.s_4 = 0;$

$p.s_5 = 0;$

$p.s_6 = 0;$

$m = \text{gillespie}(p, 1000, [< p.\text{Time} >, < p.s_4 >]);$

$\text{graph}(m);$

From each run of the simulation we obtain a graph of the time evolution of number of molecules of the protein dimer. Five such graphs are shown in Fig 7.2-7.6:

We can observe from these 5 graphs that the number of molecules of the protein dimer first increase from 0 to 25 during the time interval $[0, 2000]$, then oscillates around 23 as time goes on. This coincides with our earlier discussion about the
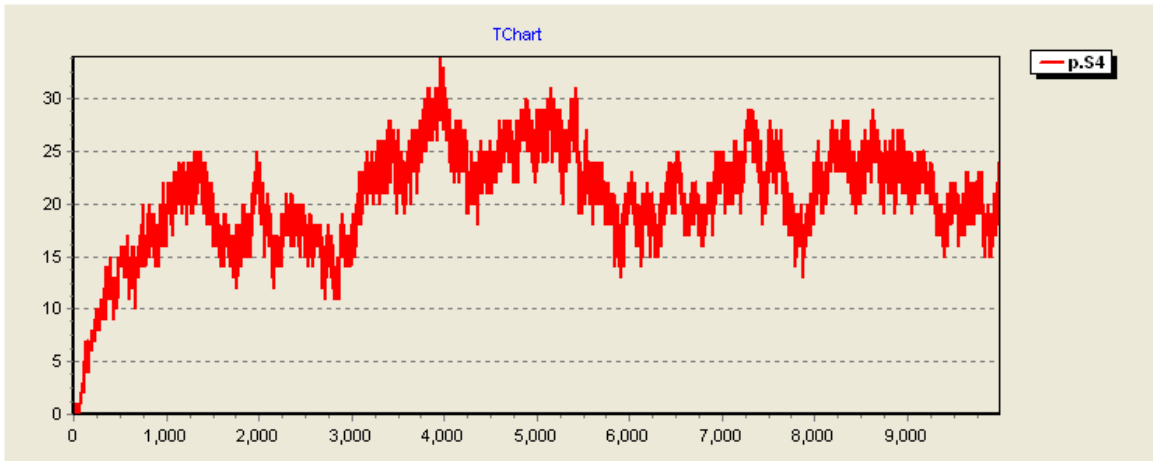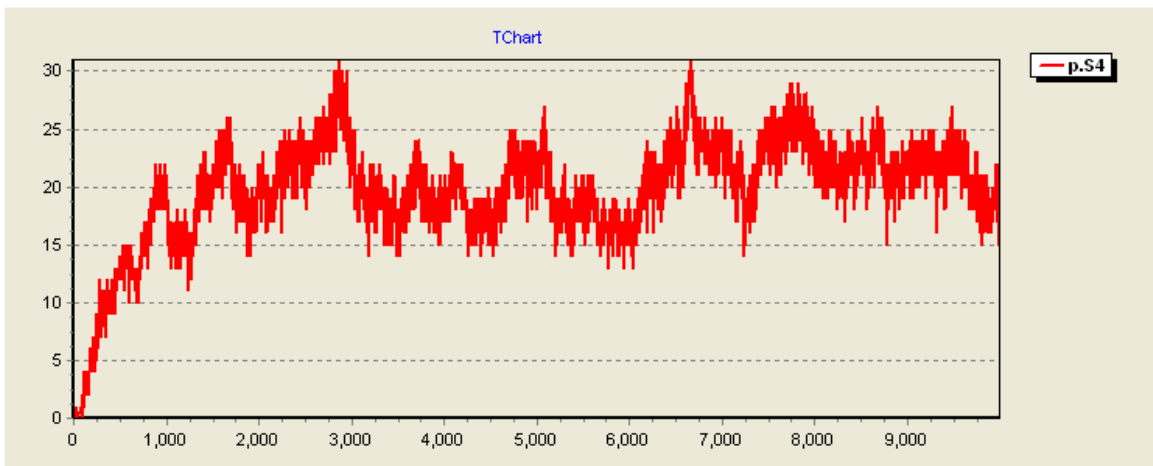
Figure 7.2: Outcome of Simulation 1.

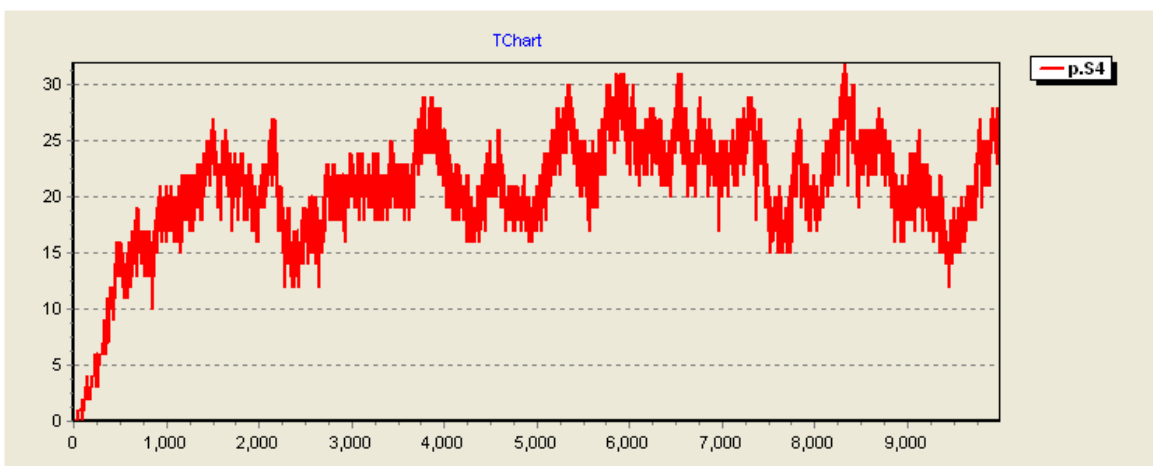Figure 7.3: Outcome of Simulation 2.
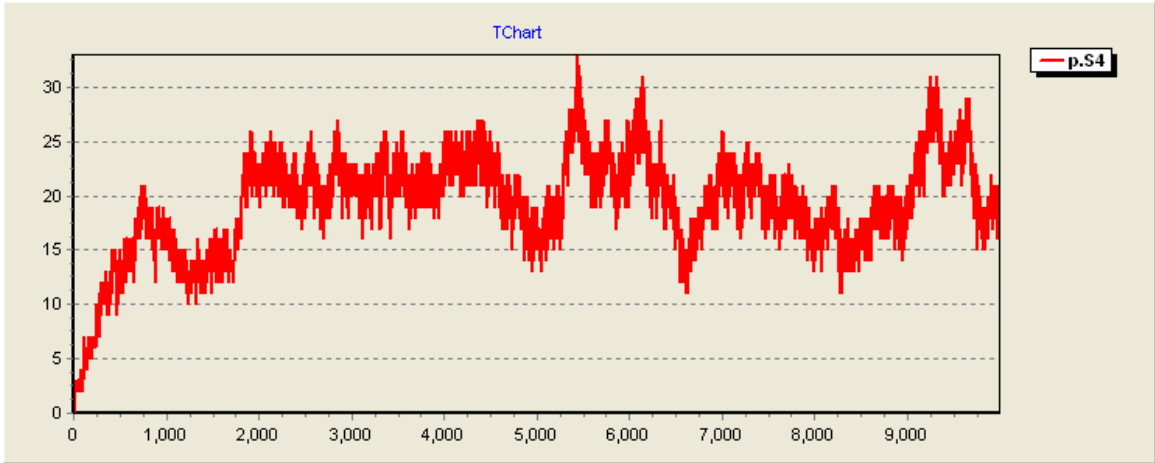

Figure 7.4: Outcome of Simulation 3.
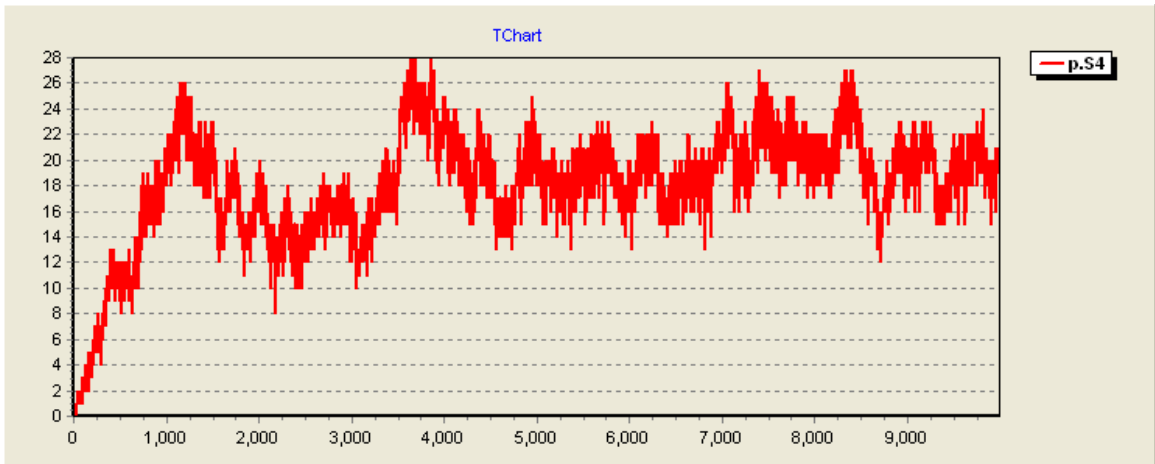
Figure 7.5: Outcome of Simulation 4.

Figure 7.6: Outcome of Simulation 5.

theorem of the long time behavior of the Master equation: as time increases, the probability of a particular state of the system approaches to a constant, which is determined by the eigenvalue 0 of the matrix A in the Master equation.

Bibliography

[1] Adam Arkin, John Ross, Harley H. McAdams, *Stochastic Kinetic Analysis of Developmental Pathway Bifurcation in Phage λ-Infected Escherichia Coli Cells*, Genetics **149**, 1633-1648, 1998

[2] Carl Branden, John Tooze, *Introduction to Protein Structure*, Garland Publishing, Inc., New York, 1991

[3] James M. Bower, Hamid Bolouri, *Computational Modeling of Genetic and Biochemical Networks*, The MIT Press, Cambridge, 2001

[4] Neil A. Campbell, Jane B. Reece, Lawrence G. Mitchell, *Biology*, Addison Wesley Longman, Inc., Menlo Park, 1999

[5] Yang Cao, Daniel T. Gillespie, Linda R. Petzold, *The Slow-Scale Stochastic Simulation Algorithm*, 2004

[6] P. Erdi, J. Toth, *Mathematical Models of Chemical Reactions: Theory and applications of deterministic and stochastic models*, Princeton University Press, Princeton, 1989

[7] C. W. Gardiner, *Handbook of Stochastic Methods*, Springer-Verlag, Berlin, 1996

[8] Daniel T. Gillespie, *A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions*, Journal of Computational Physics **22**, 403-434, 1976

[9] Daniel T. Gillespie, *Exact Stochastic Simulation of Coupled Chemical Reactions*, Journal of Physical Chemistry **81**, 2340-2361, 1977

[10] Sidney Golden, *Quantum Statistical Foundations of Chemical Kinetics*, Oxford Mathematical Monographs, 1969

[11] Peter J. E. Goss, Jean Peccoud, *Quantitative Modeling of Stochastic Systems in Molecular Biology by Using Stochastic Petri Nets*, Proc. Natl. Acad. Sci. USA **95**, 6750-6755, 1998

[12] E. Hewitt, K. Stromberg, *Real and Abstract Analysis*, Springer, New York, 1975

[13] Hidde De Jong, *Modeling and Simulation of Genetic Regulatory Systems: A literature Review*, Journal of Computational Biology **9**, 67-103, 2002

[14] Samuel Karlin, Howard E. Taylor, *A First Course in Stochastic Processes*, Academic Press, Burlington, 1966

[15] Thomas G.Kurtz, *The Relationship between Stochastic and Deterministic Models for Chemical Reactions*, J. Chem. Phys **57**, 2976-2978, 1972

[16] HARLEY H.McAdams, Adam Arkin, *Stochastic Mechanisms in Gene Expression*, Proc. Natl. Acad. Sci. USA **94**,814-819, 1997

[17] Wolfgang Reisig, Grzegovz Rozenberg, *Lectures on Petri Nets I: Basic Models*, Springer-Verlag, Berlin, 1998

[18] Sidney I. Resnick, *Adventures in Stochastic Processes*, Birkhäuser Boston, Boston, 1992

[19] Madeline A.Shea, Gary K.Ackers, *The O_R Control System of Bacteriophage Lambda; A Physical-Chemical Model for Gene Regulation*, J. Mol. Biol **181**, 211-230, 1985

[20] Paul Sjöberg, *Numerical Solution of the Master Equation in Molecular Biology* , Master Thesis, http://user.it.uu.se/ pauls/masterreport.pdf, 2002

[21] Peter Walters, *An Introduction to Ergodic Theory; Theorem 0.16*, Springer-Verlag, New York, 1982

[22] *http://www.sys-bio.org/Jarnac.htm.*