Modeling of Stochastic Non-Linear Chemical Reaction Networks

A THESIS
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Michail Vlysidis

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advised by Prodromos Daoutidis

June, 2018

Acknowledgments

The journey of pursuing a doctorate degree is a challenging one with a lot of highs and lows. I am grateful that I was surrounded by people who made the highs more enjoyable and the lows less painful and that I found the power to overcome all the obstacles and challenges. During the past five years, I became a completely different person in every aspect of my life and I am fortunate enough to receive a lot of support and love.

I want to start by thanking my adviser, Professor Prodromos Daoutidis, for his support and all the useful conversations about academia and life in general. His door was always open for me especially during the last year. I also want to thank Dr. Yiannis Kaznessis for his collaboration the past five years. Yiannis gave me the opportunity to work with stochastic kinetics and he was always supportive and excited about my ideas. He made me believe in my ideas and helped me grow as a researcher.

I am grateful to Professor Daniel Frisbie for his unconditional support the last year of my studies. He made the ending of my studies significantly less stressful. I also acknowledge Professors David Morse and Matthew Neurock for their willingness to serve as my committee members and for the insightful conversations and brilliant comments.

I acknowledge the help of my current and past group mates, especially Dr. Brittany and Dr. Giota for all of their academic and personal support. I want to thank my past and current roommates at Greek House: George, Josh, Tommy, Peter and Sanket, for helping me adapt in Minnesota. I am grateful to the Greek student community here, especially

Panagiotis, Charilaos and Giorgos, for making Minnesota feel a bit like home. Special thanks to Emily and my close friend in Greece, Harrys, for making me feel special and adding an enjoyable spice to my life. My deep gratitude to Dr. Petros Vassilakos for going above and beyond being a great friend; Petros had one of the most important contributions to improving my life the past five years.

Last but not least, a special thanks to my family: Antonis, Stauroula, Panagiota, Andromachi, Eftichia, Kosta and Eugenia, for their unconditional love and support. The weekend morning video calls helped miss Greece to a lesser degree. We went on this journey together, even though we were in different parts of the world.

Throughout my academic experience here, I have evolved as a human being and a critical thinker and I honestly believe this past five years equipped me with the necessary mindset to have a more enjoyable life. Thank you all for making this journey a happy and memorable one.

To my parents, Αντώνη and Σταυρούλα, and my sister, Παναγιώτα,
for their unconditional love and support

Abstract

Biological systems are wonderfully complex. In order to gain a better understanding on how the complexity dictates the biological functions and gives rise to different cellular phenotypes, it is important to investigate the underlying dynamic interactions of the biomolecular components. Traditionally, these system dynamics have been described using continuous deterministic mathematical models. However, small biomolecular systems are inherently stochastic. Indeed, fluctuations of molecular species are substantial in living organisms and may result in significant variation in cellular phenotypes. Since stochastic models view molecules as discrete chemical entities and the chemical events as random processes it is possible to better investigate the variability of cellular processes in detail and to gain fundamental insight.

Event though, stochastic reaction models are useful tools to describe the behavior of many biological systems, there remains a lack of numerical methods available to calculate the stationary probability distributions of stochastic reaction networks. The Chemical Master Equation (CME), which is the most detailed mathematical model that can describe stochastic behaviors, has minimal contribution due to its complexity. This work emphasizes on solving CME by a set of differential equations of probability moments, called moment equations; especially by employing Zero-Information closure scheme (ZI closure scheme). ZI closure scheme calculates the stationary probability distribution of stochastic biochemical reaction networks. The method postulates that only a finite number of probability moments is necessary to capture all of the system's information, which can be achieved by maximizing the information entropy of the system.

The work presented herein focuses on improving ZI closure scheme and resolve its numerical difficulties. The updated ZI closure scheme is a powerful tool for calculating stationary probability distribution for a variety of stochastic networks. Furthermore, we use the updated scheme to explore why deterministic and stochastic models of chemical reaction kinetics can give starkly different results when the deterministic model exhibits more than one stable solution. The closure scheme makes it possible to investigate the behavior of a system for a wide variety of volumes. For the first time, results for the mesoscopic region of stochastic networks are obtained with a numerical method.

This work also focuses on creating approaches for calculating the time transient behavior of stochastic reaction networks, since most of the numerical methods focus only on stationary probability distribution. We first tackle the issue by linearizing the moment equations and calculating the Jacobian matrix around the stationary probability distribution. The calculations are accurate and significantly more efficient than stochastic simulation algorithms based on Gillespie's algorithms. However, the approach fails to predict oscillatory behaviors. To further improve this approach, we present a new alternative method based on equations that depend only on Lagrange multipliers. The Lagrange multiplier equations approach is able to calculate the time-transient behavior of stochastic reaction networks, based on the maximum entropy principle. This work is an innovative step towards solving stochastic networks for different time behaviors.

Contents

List of Tables

List of Figures

Chapter 1

Introduction *

1.1 Stochastic Chemical Kinetics and Simulations

The classic kinetic approach to describe chemically reacting systems involves the use of deterministic rate laws [2]. This viewpoint considers the reactions to take place as a continuum. Given an initial state, the system will always evolve to the same final state predicted by the mathematical model [3]. However, this approach fails to take into account the fact that molecular populations are finite and countable. Rigorously, they must be described by integer variables that can only change stochastically and by discrete amounts. When dealing with large systems, at the thermodynamic limit, the classical approach is well suited and correctly describes their physical behavior, because molecular fluctuations are small and can be neglected [4]. Nevertheless, for small systems, which includes most biological chemical reaction networks, another mathematical approach must be used.

In the 1960s, McQuarrie and Oppenheim [5–7] were the first to propose a probabilis-

*Part of this discussion is adapted with permission from Pedro H. Constantino, Michail Vlysidis, Patrick Smadbeck and Yiannis N. Kaznessis, *Journal of Physics D: Applied Physics*, **49** 093001, 2016 [1]. Copyright ©2016 IOP Publishing Ltd.

tic approach to the problem and they introduced some of the theoretical concepts of what became famously known as the Chemical Master Equation (CME). A meticulous derivation of the CME, however, came only 25 years later through the work of Gillespie [8], using statistical thermodynamics and probability theory.

1.1.1 Chemical Master Equation

Consider, for example, an isothermal chemical reaction network at constant volume involving N chemical species connected through M reactions. Let $X_i(t)$ be the number of molecules of species i at a given time t, so that the vector $\boldsymbol{X}(t)$ describes the state of the system at time t. Assume the system is well stirred and the majority of the molecular collisions are nonreactive, i.e. they are just elastic collisions. In this case, it is not necessary to worry about molecular positions and velocities [8]. This means that the system can be fully described by vector $\boldsymbol{X}(t)$ at time t. If each reaction in the network is supposed elementary, a state-change matrix ν of size N x M can be defined where the elements ν_{ij} represent the stoichiometric coefficients of species i in reaction j. Then, the Chemical Master Equation (CME) is formally written as:

$$\frac{\partial P(\boldsymbol{X},t)}{\partial t} = \sum_{j=1}^{M} \left[\alpha_j\left(\boldsymbol{X}-\boldsymbol{\nu}_j\right) P\left(\boldsymbol{X}-\boldsymbol{\nu}_j,t\right)\right] - \sum_{j=1}^{M} \left[\alpha_j\left(\boldsymbol{X}\right) P\left(\boldsymbol{X},t\right)\right] \tag{1.1}$$

where, $\boldsymbol{\nu}_j$ is the j^{th} column of matrix ν. The function $\alpha_j\left(\boldsymbol{X}\right)$ is called the propensity and its physical significance is that $\alpha_j\left(\boldsymbol{X}\right) dt$ expresses the probability that one reaction j will occur in the next infinitesimal time interval $[t, t+dt]$. The form of the propensity function is given by the order of the reaction, the number of possible combinations among the reacting molecules and the reaction rate constant according to molecular physics [4].

The Chemical Master Equation governs how the probability distribution of the state space (molecular populations) changes over time. It can be simply viewed as a balance equation based on the law of conservation of probability for each possible state of the system. The two terms on the right hand side of equation (1.1) account for the rate of probability of arriving into state $\boldsymbol{X}(t)$ minus the rate of probability of leaving that state.

At a first glance, it might be hard to recognize the difficulty of solving the CME. The model is incredibly simple and extremely powerful. Nonetheless, as pointed out by Gillespie [4], for each possible combination of reactant molecules there will be one extra equation, which makes the CME to be in fact a system of coupled differential equations that can only be solved analytically for very simple problems. As the reaction network becomes larger and the nonlinearity of the problem increases, even numerical solutions become challenging to obtain, if not impossible.

1.1.2 The Stochastic Simulation Algorithm (SSA) and Variations

In 1976, Gillespie [9] proposed that instead of solving the CME for the probability distribution, it would be simpler to simulate individual reaction trajectories of $\boldsymbol{X}(t)$ over time. Evidently, this is not the same as suggesting a numerical solution to the master equation. Yet, by making use of the ergodic hypothesis, if one is able to generate a large ensemble of trajectories and sample the state space, it is possible to reconstruct the probability distribution as an exact solution by statistical analysis.

For that reason, Gillespie [10] developed the Stochastic Simulation Algorithm (SSA), which employs a kinetic Monte Carlo procedure to generate time trajectories. The main idea is to select which reaction will take place (integer random variable) and the time for the next reaction (exponential random variable) by drawing two uniformly distributed random numbers and applying consistent mathematical criteria. Essentially, given an initial condition, the system transitions from one state to another through jump Markov processes, which have a memoryless property, and thus evolves in time stochastically. The major issue with this approach, however, is the computational cost. Because every reaction event is simulated and many trajectories are needed in order to obtain a good solution, the SSA becomes considerably slow for most systems of interest. This is especially true if the system to be dealt with involves multiple time-scale phenomena (stiff systems).

Many other suggestions have been made by the scientific community with the purpose of improving the SSA and its efficiency. Gibson & Bruck [11], for instance, proposed the Next Reaction method, which creates a priority queue for reactions to happen and is able to use only one uniform random number per event, thus saving computational time. Other similar works include that of Cao & Petzold [12], Lok & Brent [13] and McCollum *et al*. [14]. However, as long as these methods are still simulating every single reaction event through step by step, the process remains too slow, even though it might be very efficient. It becomes, then, necessary to make use of approximate solutions in order to gain speed in computation.

1.1.3 Multi-scale Hybrid Models for Simulations

Most chemical reaction networks nonetheless span multiple time scales and different population sizes for each species. The reaction events can be rare or frequent, while the species populations can be continuous or discrete. Therefore, in the last two decades, much attention has been given to the combination of modeling formalisms through the partition of the problem space into more homogeneous regions and the creation of hybrid methods. Virtually, five different regimes for multi-scale simulations of chemical reactions can be identified according to time and population size: slow-discrete, slow-continuous, fast-discrete, fast-continuous stochastic and fast-continuous deterministic [15].

The essential idea of hybrid methods is to use approximate solutions to describe the dynamics of fast reactions, while applying exact solutions for slow ones. For instance, one of the simplest strategies was developed by Rao & Arkin [16] assuming quasi-steady state for fast reactions and applying the regular SSA for slow events. Another approach by Puchałka & Kierzek [17] used a next reaction variant for the slow regions and tau-leaping with Poisson distribution for fast regions. Haseltine & Rawlings [18] modeled a discrete jump Markov chain for the slow reactions and approximated the fast region by a continuous Markov process. Finally, Kaznessis *et al*. [19–21] developed methods to model fast reactions through continuous Markov chains, while slow reactions are described by a slight modification of the Next Reaction method proposed by Gibson & Bruck [11].

Even though hybrid methods are reliable and faster alternatives of SSA, they are still computationally expensive and less accurate than SSA.

1.2 Moment Equations of Stochastic Reaction Networks

It still exists the need to develop more efficient ways to solve stochastic systems. One popular alternative among researchers to explain such systems is the use of time derivatives of probability moments.

Moments are expected values which are used to describe a probability distribution [22]. The first moment is related to the mean of the distribution, the second moment is related to the variance, the third to the skewness, the fourth to the kurtosis etc. [23]. In the most general case, a moment μ_i is connected to probability distribution through the following expression:

$$\mu_i = \sum_{\Omega} f_{\mu_i}(\boldsymbol{X})P(\boldsymbol{X}) \tag{1.2}$$

where f_{μ_i} is the functional form of the i^{th} moment μ_i and Ω corresponds to the N-dimensional state space. This notation still assumes that the system involves N chemical species, the system can be fully described by vector $\boldsymbol{X}$ and P is the probability distribution.

Different function for f_{μ_i} generate different moment bases with the most common ones to be the central [24], polynomial [25], jump [23] and factorial [22] moments. The CME can be correlated to a set of ODEs called moment equations involving their time derivatives:

$$\frac{\partial \boldsymbol{\mu}}{\partial t} = A\boldsymbol{\mu} + A'\boldsymbol{\mu'} + \boldsymbol{\mu_c} \tag{1.3}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\mu'}$ are vectors containing lower-order moments (moments up to a specific order) and higher-order moments, respectively. Vector $\boldsymbol{\mu}$ does not include the zero$^{\text{th}}$ moment, which is always equal to 1. $\boldsymbol{\mu_c}$ is a vector of constants and t represents time. A and A' are constant matrices that represent the linear and non-linear components of the network, respectively. In most cases, vectors $\boldsymbol{\mu}$ and $\boldsymbol{\mu'}$ have different dimensions [25]. In non-linear reaction networks the lower-order moments depend on the higher-order ones (A' is non zero), which creates difficulties in solving the system. All the different types of moments are valid for describing a probability distribution [22]. As a result, the moment equations may have different forms of expression.

In 1949 [26], Moyal was one of the first who used time derivatives of expected values in order to express stochastic systems. However, the first publication suggesting the use of probability moments for solving a stochastic chemical reaction network was from McQuarrie, Jachimowski and Russell in 1964 [27]. In their publication, they worked with non-linear reaction networks, however they limited their discussion to up to two moments. A common assumption among researchers was that up to two moments would have been enough to describe a distribution [23]. Nowadays, we know that a system can require many more moments up to any specific order; this can be even eight or ten moments [23, 28]. Since then, there have been numerous attempts to create and solve moments equations [22, 23, 25, 29]. Most of the scientific efforts are focused in two areas: 1) how to generate moment equations more efficiently and 2) how to express the higher-order moments in terms of the lower-order ones.

1.2.1 Derivation of Moment Equations

Sotiropoulos and Kaznessis [23] suggested the use of jump moments for the derivation of moment equations. Their formulation provided an analytical way to calculate the moment time derivative equations, however the approach is not computationally friendly. The analytical form derived is not recursive, which implies that the computational cost increases disproportionately with an increase of number of moments and components (combinatorial explosion) [22]. Another work that should be noted is that of C. S. Gillespie in 2009 [25]. Using a univariate moment generating function and the polynomial moments, Gillespie managed to produce an efficient way of reconstructing the moment equation. The paper also suggests the use of probability cumulants instead of probability moments in order to solve the set of equations. Probability cumulants are the natural logarithm of the moments of the generating function.

Even though the use of polynomial moments gave good results for Gillespie [25], at least for small systems, Smadbeck and Kaznessis [22] proposed a more computationally efficient approach. They integrated the use of factorial moments for the generation of moment matrices and moment equations. Through the use of factorial moments, the matrices produced have a banded form, which requires less memory to store them. Besides, factorial moment matrices can be produced faster than polynomial ones [22]. Factorial and polynomial moments can be mapped into each other and produce equivalent valid versions of moment equations [1]; they are just a different basis set [22, 30].

The rest of the subsection is denoted to the derivation of moment equations with factorial moment base. The i^{th} factorial moment can be denoted as $\mu_i = \{X_j^i\}$ and its functional form is:

$$f_{\mu_i} = f_{\{X_j^i\}} = \sum_{X_j=0}^{\infty} \frac{X_j!}{(X_j - i)!} \tag{1.4}$$

Thus the general form of factorial moments is given by the expression:

$$\{X_1^{m_1} \cdot X_2^{m_2} \cdot ... \cdot X_N^{m_N}\} =$$
$$\sum_{X_1=0}^{\infty} \sum_{X_2=0}^{\infty} ... \sum_{X_N=0}^{\infty} \frac{X_1!}{(X_1 - m_1)!} \frac{X_2!}{(x_2 - m_2)!} ... \frac{X_N!}{(x_N - m_N)!} P(\boldsymbol{X}, t) \tag{1.5}$$

with the above moment being of order $m = m_i + m_j + ... + m_N$. One of the reasons for using factorial moments is that they are the space derivatives of Z-transform of the probability distribution, which is simply a spatial transformation having the following form:

$$G(\boldsymbol{S}, t) = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} ... \sum_{x_n=0}^{\infty} s_1^{x_1} s_2^{x_2} ... s_n^{x_n} P(\boldsymbol{X}, t) \tag{1.6}$$

where $\boldsymbol{S} = [s_1, s_2, ..., s_n]^T$ is the continuous transformation of $\boldsymbol{X}$, with s_i representing the space variable of component i. Differentiating equation (1.6) with respect to space yields:

$$\left. \frac{\partial G(\boldsymbol{S}, t)}{\partial s_i} \right|_{\boldsymbol{S}=\mathbf{1}} = \{x_i\} \tag{1.7}$$

$$\left. \frac{\partial^2 G(\boldsymbol{S}, t)}{\partial s_i^2} \right|_{\boldsymbol{S}=\mathbf{1}} = \{x_i^2\} \tag{1.8}$$

Besides, it is also true that: $G(\boldsymbol{S}, t)|_{\boldsymbol{S}=\mathbf{1}} = 1$. Thus, it is easy to generate factorial moments through the Z-transformation of the probability distribution. Differentiating the function $G(\boldsymbol{S}, t)$ with respect to time:

$$\frac{\partial G(\boldsymbol{S}, t)}{\partial t} = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} ... \sum_{x_n=0}^{\infty} s_1^{x_1} s_2^{x_2} ... s_n^{x_n} \frac{\partial P(\boldsymbol{X}, t)}{\partial t} \tag{1.9}$$

Applying it to the CME (1.1) with some mathematical manipulations results into:

$$\frac{\partial G(\boldsymbol{S},t)}{\partial t} = \sum_{\mu} \frac{k_\mu}{\prod_{i=1}^{n} \nu_{i,\mu}^{+}!} \left[\prod_{i=1}^{n} s_{i,\mu}^{\nu_{i,\mu}^{-}} - \prod_{i=1}^{n} s_{i,\mu}^{\nu_{i,\mu}^{+}} \right] \frac{\partial^{\sum_{i=1}^{n} \nu_{i,\mu}^{+}} G(\boldsymbol{S},t)}{\partial \prod_{i=1}^{n} s_{i,\mu}^{\nu_{i,\mu}^{+}}} \tag{1.10}$$

where the $\nu_{i,\mu}^{+}$ and $\nu_{i,\mu}^{-}$ refers to the absolute values of the stochiometries of the reactants and products, respectively [22]. Equation, 1.10 is the Z-transformation of the Chemical Master Equation (Z-CME). Compared to the CME, that depends on the discrete space variable $\boldsymbol{X}$, the Z-CME depends on a continuous space variables.

Combining equations (1.7), (1.8) and (1.10), one can construct the first two moment equations for any chemical reaction network. Hence, by differentiating equation (1.10) with respect to the appropriate space variable, matrices A and A' we can be produced from equation (1.3).

So, the moment equations can be obtained through the Z-transformation of CME. Even though, the necessary number of extra higher-order moments may be determined, it is still an open research interest to find what is the minimum number of lower-order moments needed to describe a given reaction network [22].

1.2.2 Closure Schemes Approximations

Although the idea of using moments to describe the dynamics of stochastic reaction networks is not a new concept [5], only recent advances in mathematical and numerical formalisms have made this idea fruitful. Numerous algorithms have been developed to efficiently produce moment equations [22, 25, 31–33]. However, there is a lack of numerical methods to solve moment equations for an arbitrary reaction network. The difficulty stems from the dependence of lower-order moments on the higher-order ones (as demonstrated in equation (1.3)). Most scientific effort is focused on solving issues with the non-linearity of moment equations through closure schemes, which are mathematical techniques that relate higher and lower-order moments. Despite these efforts, there is not a solution that has broad system applicability.

The most common approach in the literature to tackle this issue and close the scheme is to assume a specific form of the species behavior. By assuming a specific component's behavior one can relate the higher-order moments to the lower-order ones [34–36]. Linear noise approximation (LNA), 2MA and 3MA are well-known approaches to produce analytical expressions for the first lower-order moments. Even though these techniques generate accurate results for linear systems, they lack the ability to correctly predict non-linear behaviors [31,33,36,37]. An important subset of the literature approaches the issue by assuming a specific form of the system's distribution, such as normal, poison, lognormal, gamma etc. [32,33,35,38]. Singh and Hespanha [29] reported some of the common approaches in literature. In cases that the observed components behave in a known fashion, such distribution assumptions produce reasonably accurate results. However, in most systems, there is a lack of knowledge of the species development and thus such assumptions breakdown or become impractical.

A reasonable broad approach comes from the observation that a more appropriate way of describing something unknown is by choosing the most unbiased estimation. Therefore, the description should be based on the distribution that is most uncertain. Smadbeck and Kaznessis [28] suggested the use of information entropy in order to closure the scheme, i.e. to the connect the higher-order moments with lower-order ones.

1.3 Thesis Scope and Organization

The focus of this work is to improve the maximum entropy approach for solving stationary probability distributions of stochastic reactions networks and create new numerical approaches on calculating the time transient behavior of such systems.

Chapter 2 presents in exhaustive detail the algorithmic steps of the Zero-Information (ZI) closure scheme. ZI closure scheme is a method for solving the Chemical Master Equation and moment equations of stochastic chemical reaction networks based on the maximum entropy principle. This chapter is a modified version of [39] with some parts from [1] with permission. The objective of this chapter is two-fold: first to present the algorithm with sufficient didactic value for a non-expert, yet patient and self-motivated reader to confidently reconstruct the algorithm with any available programming tools, including the ubiquitous paper and pencil. Second, this chapter lays bare the theoretical and numerical underpinnings of ZI closure for the expert audience to confidently criticize and propose new, improved implementation solutions. The detailed algorithm of this approach can be found in Chapter 7.

In order to make ZI closure scheme algorithm broadly available, we have created ZICS, a Windows application employing the scheme. Chapter 3 provides detailed steps on how to use the application. The application calculates the stationary probability distribution for a user-defined reaction network with minimal computational cost. ZICS accurately produces results for multidimensional, multistable, nonlinear reaction networks. Additionally, it provides an algorithm that generates factorial moment equations for reaction networks, as well as a means for visualizing the probability distribution of molecular components. The application is compatible with Matlab, version 2016b and later, so it can be also used on non-Windows machines.

Using this tool, in Chapter 4 we set out to explore why deterministic and stochastic models of chemical reaction kinetics can give starkly different results when the deterministic model exhibits more than one stable solution. In one example, the bimodal probability distribution exhibited by the stochastic Schlögl model gradually collapses to a unimodal distribution for mesoscopic systems, even for kinetic constant values that result in two stable attractors in the deterministic model. Using ZI closure scheme, probability distributions for large system sizes that were previously unattainable by stochastic simulation algorithms are now computed. We observe and quantify the paradoxical discrepancies between stochastic and deterministic models and explain this behavior by postulating maximum entropy non-equilibrium steady states.

Even though ZI closure scheme is a useful tool for calculating stationary probability distributions, it is unable to provide insight about time dynamics. Based on this moment closure steady state approach, in Chapter 5 we propose a linearization method for probability moment equations to calculate the time-transient behavior of stochastic reaction networks close to steady state. This new method can be used to model reaction networks away from the thermodynamic limit, including biomolecular and metabolic reaction networks in biological systems.

The theoretical underpinnings of the linearization method are presented in Chapter 5. This chapter is a reprint with permission from [40]. The approach linearizes moment equations by calculating the Jacobian matrix around the stationary probability distribution. In order to demonstrate the method, we present examples of stochastic reaction networks and compute their dynamic behavior.

The linearization technique require significant assumptions in order to solve stochastic reaction networks. Chapter 6 presents a new alternative method based on equations that depend only on Lagrange multipliers. The Lagrange multiplier equations approach is able to calculate the time-transient behavior of stochastic reaction networks, based on the maximum entropy principle. We report detailed steps on how to transform any kind of moment equations to Lagrange multiplier ones. In order to demonstrate the method, we present examples of non-linear stochastic reaction networks with varying complexities, including multistable and oscillatory systems. We find that the new approach is as accurate and significantly more efficient than stochastic simulation algorithms based on Gillespie's algorithms. This work is an innovative step towards solving stochastic networks for different time behaviors. The detailed algorithm of this approach is reported in Chapter 7.

Finally, Chapter 8 presents the concluding remarks and future directions of this work. Some of the reprinted chapters have been modified to improve consistency throughout the document.

Chapter 2

ZI Closure Scheme: A Method to Solve and Study Stochastic Reaction Networks *

This chapter lays the fundamentals of a proposed updated Zero-Information (ZI) closure scheme approach. The original ZI closure scheme was created by Smadbeck & Kaznessis [28]. The objective of this chapter is to present in exhaustive detail the algorithmic steps of the updated method. The method solves the moment equations of stochastic chemical reaction networks.

The original algorithm was employing the null space of the system. However, the calculation of the null space for reaction networks with more than one degrees of freedom is numerically impractical. The proposed approach of this chapter and the algorithm of Chapter 7 bypasses this issue while maintaining the same principles and accuracy. Additionally, the original approach [28] was considering the zero$^{\text{th}}$ moment values as an uknown and thus the zero$^{\text{th}}$ Lagrange multiplier independent. These assumptions can

*Reprinted (adapted) with permission from Michail Vlysidis, Pedro H. Constantino and Yiannis N. Kaznessis, *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology. Springer, Cham*, 2017 [39]. Copyright ©Springer International Publishing AG 2017. Part of this discussion is also adapted with permission from Pedro H. Constantino, Michail Vlysidis, Patrick Smadbeck and Yiannis N. Kaznessis, *Journal of Physics D: Applied Physics*, **49** 093001, 2016 [1]. Copyright ©2016 IOP Publishing Ltd.

cause some important numerical instabilities and provide unphysical results. The approach of this chapter has corrected the aforementioned issues by explicitly recognizing the dependence of the zero$^{\text{th}}$ Lagrange multiplier.

2.1 Introduction

In 1948, C. E. Shannon published a simple and revolutionary paper in The Bell System Technical Journal [41], where he introduced the concept of information entropy (represented as H) as a measure of uncertainty or unexpectedness. The idea was that the more unexpected an event was, the more new information it carried within itself. If, on the other hand, an outcome of a system can be inferred from previous knowledge, then the new event is redundant, as it provides little new information. It has been said that when Shannon asked for advice to the famous mathematician Newmann about how to call his new measure, he replied: ”You call it entropy, because most people do not know what entropy really is, and if you use the word ’entropy’ in an argument, you will win every time!” [42] It has also been suggested that what Shannon did for information theory was equivalent to ”what Gibbs did for physical chemistry” [43].

However, it was not until 1957 that E. T. Jaynes first proposed the Maximum Entropy Principle (MEP), applying it in the context of statistical mechanics and drawing an analogy with the second law of thermodynamics [44]. If some information about a probability distribution is known *a priori*, then the MEP states that one should look for the distribution with maximum uncertainty, while satisfying the given information (unbiased estimate).

Based on those principles, we have developed the ZI closure scheme for the system of moment equations. It is assumed that all information necessary to build the exact probability distribution described by the Chemical Master Equation (CME) is contained within a finite number of lower-order moments. In that case, all higher-order moments add no information to the problem and can, therefore, be obtained from the maximum entropy probability distribution.

Using Shannon's definition, for a discrete multivariate system of N components, the entropy is given by:

$$H = -\sum_{\Omega} P(\boldsymbol{X}) \ln P(\boldsymbol{X}) \tag{2.1}$$

where Ω corresponds to the N-dimensional state space of all possible values for $\boldsymbol{X} = (X_1 ... X_N)$, used here to simplify the notation of multiple summation signs. $P(\boldsymbol{X})$ represents the discrete multivariate probability distribution.

Assuming that the first M lower-order moments are known, the entropy defined in equation (2.1) must be maximized with respect to these constrains. This can be accomplished through the method of Lagrange multipliers:

$$\Lambda = H - \sum_{j=0}^{M} \lambda_j \left[\mu_j - \sum_{\Omega} f_{\mu_j}(\boldsymbol{X}) P(\boldsymbol{X}) \right] \tag{2.2}$$

where Λ is the Lagrangian and λ_j is the Lagrange multiplier associated with the lower-order moment μ_j. The function f_{μ_j} defines the j-th lower moment, as discussed in Chapter 1. Knowing the lower-order moments, their Lagrange multipliers might be computed using a root-finder numerical method such as Newton-Raphson. Hence, the maximum information entropy distribution is written as follow [45]:

$$P_H(\boldsymbol{X}) = \exp \left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right] \tag{2.3}$$

The subscript H emphasizes the maximum entropy distribution. Now with the most unbiased estimate for the distribution being known, the higher-order moments can be calculated as:

$$\mu_i' = \sum_{\Omega} f_{\mu_i'}(\boldsymbol{X}) P_H(\boldsymbol{X}) \tag{2.4}$$

Replacing the originally unknown higher moments with the ones calculated assuming maximum uncertainty and zero information contribution, the system of differential equations for lower moments is now closed and ready to be integrated:

$$\frac{\partial \boldsymbol{\mu}}{\partial t} = A\boldsymbol{\mu} + A'\boldsymbol{\mu'} + \boldsymbol{\mu_c} \tag{2.5}$$

As discussed in Chapter 1, $\boldsymbol{\mu}$ and $\boldsymbol{\mu'}$ are the lower and higher-order moment vectors, $\boldsymbol{\mu_c}$ is a constant vector and A & A' are constant matrices. If both lower and higher-moments are expressed in terms of the maximum entropy distribution and the left-hand side of moment equations (equation 2.5) is set to zero, then it becomes a system of M nonlinear algebraic equations with M unknown Lagrange multipliers:

$$0 = A\boldsymbol{\mu} + A'\boldsymbol{\mu'} + \boldsymbol{\mu_c} \tag{2.6}$$

Again, using a multivariable root-finder method such as Newton-Raphson, determining the solution to this system will specify the stationary probability distribution.

In the rest of the chapter, the ZI closure scheme for the system of moment equations is presented in full mathematical detail. The method is applied to a specific reaction network, the Schlögl model [46], as an illustrative study case. The Schlögl model is a single-component theoretical system that can exhibit bistability. The simplicity of the state space (single-component) alongside with the complexity of the probability distribution (bistability) makes the model a compelling studying example for the ZI closure scheme.

The analytical derivation of the CME of the Schlögl model is presented first. Then the derivation of the associated moment equations is discussed in details. Next, we move on to explicitly discuss the theoretical framework supporting the ZI closure scheme. Finally, some results comparing ZI closure scheme's accuracy to kinetic Monte Carlo simulations are included.

2.2 Theoretical Background

2.2.1 Derivation of Chemical Master Equation for the Schlögl Model

This subsection presents the derivation of the chemical master equation for the Schlögl model. The CME models the probability $P(X;t)$ that the system is in state X at time t. Here X represents the number of molecules of the sole reactant X.

The Schlögl model consists of the following reactions [46]:

$$3\,\mathrm{X} \xrightarrow{\mathrm{k_1}} 2\,\mathrm{X} \tag{2.7}$$

$$2\,\mathrm{X} \xrightarrow{\mathrm{k_2}} 3\,\mathrm{X} \tag{2.8}$$

$$\mathrm{X} \xrightarrow{\mathrm{k_3}} \emptyset \tag{2.9}$$

$$\emptyset \xrightarrow{\mathrm{k_4}} \mathrm{X} \tag{2.10}$$

The probability of reaction (2.7) occurring during an infinitesimal time interval $(t, t+dt)$ is $P_1(X) = k_1\,X(X-1)(X-2)\,dt$. For reaction (2.8) this probability is $P_2(X) = k_2\,X(X-1)\,dt$, for reaction (2.9) we obtain $P_3(X) = k_3\,X\,dt$ and for reaction (2.10) the probability is $P_4(X) = k_4\,dt$ [47], with P_i, $i = 1, 2, 3, 4$, being the probability of reaction i firing in isolation. In order for the process to be considered Markovian, at most one reaction is allowed per time interval dt [47].

In the Schlögl model, the only ways for the system to be at X number of molecules at time $t + dt$ are:

(a) The system was at X number of molecules at time t and no reaction occurred within $(t, t + dt)$,

(b) the system had X - 1 number of molecules at time t and either reaction (2.8) or reaction (2.10) occurred within $(t, t + dt)$,

(c) the system had $X + 1$ number of molecules at time t and either reaction (2.7) or reaction (2.9) occurred within $(t, t + dt)$.

The probability of event (a) is:

$$P(a) \;=\; P(no\,reaction\;occurs) \cdot P(system\;has\;X\;molecules\;at\;time\;t) = \\ = \;[1 - P(all\;reactions\;occur)]\;P(X;t)$$

Since all reactions are considered independent, this results in:

$$P(a) \;=\; [1 - P_1(X) - P_2(X) - P_3(X) - P_4(X)]\;P(X;t) = \\ = \left[1 - k_1\;X(X-1)(X-2)\;dt - k_2\;X(X-1)\;dt - k_3\;X\;dt - k_4\;dt\right]\;P(X;t) \tag{2.11}$$

Similarly, the probability of event (b) is:

$$P(b) \;=\; [P_2(X-1) + P_4(X-1)]\;P(X-1;t) = \\ = \left[k_2\;(X-1)(X-2)\;dt + k_4\;dt\right]\;P(X-1;t) \tag{2.12}$$

And for event (c) we have:

$$P(c) \;=\; [P_1(X+1) + P_3(X+1)]\;P(X+1;t) = \\ = \left[k_1\;(X+1)X(X-1)\;dt + k_3\;(X+1)\;dt\right]\;P(X+1;t) \tag{2.13}$$

Thus, the probability that the system is at state X at time $t + dt$ is:

$$P(X; t + dt) \;=\; P(a) \;+\; P(b) \;+\; P(c) \;=$$
$$\left[1 - k_1\, X(X-1)(X-2)\, dt - k_2\, X(X-1)\, dt - k_3\, X\, dt - k_4\, dt\right] P(X; t) -$$
$$+ \left[k_2\, (X-1)(X-2)\, dt + k_4\, dt\right] P(X-1; t) +$$
$$+ \left[k_1\, (X+1)X(X-1)\, dt + k_3\, (X+1)\, dt\right] P(X+1; t)$$

Rearranging the last expression:

$$P(X; t + dt) - P(X; t) = \Bigg\{ \left[k_2\, (X-1)(X-2)\; + k_4\,\right] P(X-1; t) +$$
$$- \left[k_1\, X(X-1)(X-2)\; + k_2\, X(X-1)\; + k_3\, X\; + k_4\right] P(X; t) +$$
$$+ \left[k_1\, (X+1)X(X-1)\; + k_3\, (X+1)\right] P(X+1; t) \Bigg\}\, dt \Longleftrightarrow$$

$$\frac{P(X; t + dt) - P(X; t)}{dt} = \left[k_2\, (X-1)(X-2)\; + k_4\,\right] P(X-1; t) -$$
$$- \left[k_1\, X(X-1)(X-2)\; + k_2\, X(X-1)\; + k_3\, X\; + k_4\right] P(X; t) +$$
$$+ \left[k_1\, (X+1)X(X-1)\; + k_3\, (X+1)\right] P(X+1; t)$$

And then by taking the limit $dt \longrightarrow 0$ we have:

$$
\begin{aligned}
\lim_{dt \to 0} \frac{P(X;t+dt) - P(X;t)}{dt} = \frac{\partial P(X;t)}{\partial t} = \\
= \left[k_2 \, (X-1)(X-2) \, + k_4 \right] P(X-1;t) - \\
- \left[k_1 \, X(X-1)(X-2) \, + k_2 \, X(X-1) \, + k_3 \, X \, + k_4 \right] P(X;t) + \\
+ \left[k_1 \, (X+1)X(X-1) \, + k_3 \, (X+1) \right] P(X+1;t)
\end{aligned} \tag{2.14}
$$

This is the CME for the Schlögl reaction model.

2.2.2 Derivation of Moment Equations from Chemical Master Equation

A proposed technique to solve the CME relies on calculating the probability moments. The approach is based on the idea that any probability distribution can be completely described by its moments. As discussed in Chapter 1, moments are expected values of functions of an independent random variable. Herein factorial moments are used, given by the expression:

$$
\{X^m\} = \sum_{x=0}^{\infty} \frac{x!}{(x-m)!} P(x,t) \tag{2.15}
$$

where x represents the possible values of X and m is the integer order of the moment.

A main reason for using factorial moments is that they are simple derivatives of a Z-transform of the probability distribution, which has the following form:

$$
G(s,t) = \sum_{x=0}^{\infty} s^x P(x,t) \tag{2.16}
$$

where s is the continuous transformation of X. Z-transform is also known as the probability-generating function.

Differentiating equation (2.16) with respect to s, at $s = 1$, yields:

$$\left.\frac{\partial G(s,t)}{\partial s}\right|_{s=1} = \{X\}, \qquad \left.\frac{\partial^2 G(s,t)}{\partial s^2}\right|_{s=1} = \{X^2\},$$
$$\left.\frac{\partial^3 G(s,t)}{\partial s^3}\right|_{s=1} = \{X^3\}, \qquad \left.\frac{\partial^4 G(s,t)}{\partial s^4}\right|_{s=1} = \{X^4\} \tag{2.17}$$

Similar equations hold for the rest of the moments (fifth-order, sixth-order, etc.). Note that it is also true that $G(s,t)|_{s=1} = 1$. Thus, it is easy to generate factorial moments through the Z-transform of the probability distribution.

Differentiating the function $G(s,t)$ with respect to time:

$$\frac{\partial G(s,t)}{\partial t} = \sum_{x=0}^{\infty} s^x \frac{\partial P(x,t)}{\partial t} \tag{2.18}$$

Applying this to the CME for the Schlögl model (2.14), one obtains the Z-transformed Chemical Master Equation (Z-CME):

$$\frac{\partial G\,(s,t)}{\partial t} =$$
$$k_1\left(s^2 - s^3\right)\frac{\partial^3 G}{\partial s^3} + k_2\left(s^3 - s^2\right)\frac{\partial^2 G}{\partial s^2} + k_3\,(1-s)\,\frac{\partial G}{\partial s} + k_4\,(s-1)\,G \tag{2.19}$$

Through Z-CME one can generate moment equations. Taking the first derivative of equation (2.19) with respect to s yields:

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{\partial G}{\partial s}\right) = \frac{\partial}{\partial s}\left(\frac{\partial G}{\partial t}\right) &= k_1\left(2s - 3s^2\right)\frac{\partial^3 G}{\partial s^3} + k_1\left(s^2 - s^3\right)\frac{\partial^4 G}{\partial s^4}\\ &+k_2\left(3s^2 - 2s\right)\frac{\partial^2 G}{\partial s^2} + k_2\left(s^3 - s^2\right)\frac{\partial^3 G}{\partial s^3} + k_3\,(-1)\,\frac{\partial G}{\partial s}\\ &+k_3\,(1-s)\,\frac{\partial^2 G}{\partial s^2} + k_4\,(1)\,G + k_4\,(s-1)\,\frac{\partial G}{\partial s}\end{aligned} \tag{2.20}$$

Setting $s = 1$:

$$\left.\frac{\partial}{\partial t}\left(\frac{\partial G}{\partial s}\right)\right|_{s=1} = -k_1\left.\frac{\partial^3 G}{\partial s^3}\right|_{s=1} + k_2\left.\frac{\partial^2 G}{\partial s^2}\right|_{s=1} - k_3\left.\frac{\partial G}{\partial s}\right|_{s=1} + k_4\left.G\right|_{s=1} \tag{2.21}$$

Or from equations (2.17):

$$\frac{\partial\{X\}}{\partial t} = -k_1\{X^3\} + k_2\{X^2\} - k_3\{X\} + k_4 \tag{2.22}$$

Equation (2.22) is the equation for the first-order factorial moment of $P(X,t)$. In order to generate the moment equation for the second moment one needs to take the second derivative of equation (2.19) with respect to s (or the derivative of equation (2.20)):

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{\partial^2 G}{\partial s^2}\right) = \frac{\partial^2}{\partial s^2}\left(\frac{\partial G}{\partial t}\right) = k_1\left(2-6s\right)\frac{\partial^3 G}{\partial s^3} + k_1\left(2s-3s^2\right)\frac{\partial^4 G}{\partial s^4} \\ +k_1\left(2s-3s^2\right)\frac{\partial^4 G}{\partial s^4} + k_1\left(s^2-s^3\right)\frac{\partial^5 G}{\partial s^5} + k_2\left(6s-2\right)\frac{\partial^2 G}{\partial s^2} \\ +k_2\left(3s^2-2s\right)\frac{\partial^3 G}{\partial s^3} + k_2\left(3s^2-2s\right)\frac{\partial^3 G}{\partial s^3} + k_2\left(s^3-s^2\right)\frac{\partial^4 G}{\partial s^4} - k_3\frac{\partial^2 G}{\partial s^2} \\ +k_3\left(-1\right)\frac{\partial^2 G}{\partial s^2} + k_3\left(1-s\right)\frac{\partial^3 G}{\partial s^3} + k_4\frac{\partial G}{\partial s} + k_4\left(1\right)\frac{\partial G}{\partial s} + k_4\left(s-1\right)\frac{\partial^2 G}{\partial s^2}\end{aligned} \tag{2.23}$$

Again, by setting $s = 1$:

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{\partial^2 G}{\partial s^2}\right)\bigg|_{s=1} = -4k_1\left.\frac{\partial^3 G}{\partial s^3}\right|_{s=1} - k_1\left.\frac{\partial^4 G}{\partial s^4}\right|_{s=1} - k_1\left.\frac{\partial^4 G}{\partial s^4}\right|_{s=1} \\ +4k_2\left.\frac{\partial^2 G}{\partial s^2}\right|_{s=1} + k_2\left.\frac{\partial^3 G}{\partial s^3}\right|_{s=1} + k_2\left.\frac{\partial^3 G}{\partial s^3}\right|_{s=1} - k_3\left.\frac{\partial^2 G}{\partial s^2}\right|_{s=1} \\ -k_3\left.\frac{\partial^2 G}{\partial s^2}\right|_{s=1} + k_4\left.\frac{\partial G}{\partial s}\right|_{s=1} + k_4\left.\frac{\partial G}{\partial s}\right|_{s=1}\end{aligned} \tag{2.24}$$

Or from equation (2.17):

$$\frac{\partial\{X^2\}}{\partial t} = -2k_1\{X^4\} + \left(-4k_1+2k_2\right)\{X^3\} + \left(4k_2-2k_3\right)\{X^2\} + 2k_4\{X\} \tag{2.25}$$

This is the second-order factorial moment equation. All higher order moments can be constructed with this recursive algorithm.

For non-linear reactions this set of equations must, in principle, be extended to infinite order. The reason becomes evident by looking at equation (2.22), where the first moment depends on the second and the third. In equation (2.25) it is evident that $\{X^2\}$ depends on $\{X^3\}$ and $\{X^4\}$. Similarly, in the equation for $\{X^M\}$, where M is an arbitrarily large, yet finite, moment, we would find terms like $\{X^{(M+1)}\}$, $\{X^{(M+2)}\}$, etc.

We can concisely write the moment equations as follows:

$$\frac{\partial \boldsymbol{\mu}}{\partial t} = A\boldsymbol{\mu} + A'\boldsymbol{\mu'} + \boldsymbol{\mu_c} \tag{2.26}$$

where $\boldsymbol{\mu}$ is the vector of lower-order moments:

$$\boldsymbol{\mu} = \left[1\ \{X\}\ \{X^2\}\ \{X^3\}\ \{X^4\}\ \{X^5\} \ldots \{X^M\}\right]^T \tag{2.27}$$

and $\boldsymbol{\mu'}$ is the vector of higher-order moments:

$$\boldsymbol{\mu'} = \left[\{X^{(M+1)}\}\ \{X^{(M+2)} \ldots\}\right]^T \tag{2.28}$$

$\boldsymbol{\mu_c}$, A and A' are constructed with coefficients from moment equations. The major challenge with solving the moment equations (equation (2.26)) is the result of non-zero elements present in matrix A'. The set never closes for non-linear reactions, requiring an infinite number of moment equations.

One insight to move past this impasse is that for reaction networks with finite state space we may anticipate a probability distribution that can be adequately described by a finite number of moments. Note this argument is void of mathematical rigor. The insight is purely based on physical intuition. Let us accept it for now with the understanding that it may or may not serve us well, depending on the reaction network.

Empirically, we find that the probability distribution of the Schlögl model can indeed be described within reasonable numerical accuracy with up to 10 lower-order moments (figure 2.2). For the sake of brevity, the algorithm is presented with the assumption that five moments suffice. The focus is only on steady state probability distributions. Non-steady state solutions are the subject of Chapters 5 & 6.

At steady state, moment equations are denoted by equation (2.6). For the Schlögl model with up to five lower-order moments (excluding the zero order moment):

$$\boldsymbol{\mu} = \left[\{X\}\ \{X^2\}\ \{X^3\}\ \{X^4\}\ \{X^5\}\right]^T$$
$$\boldsymbol{\mu'} = \left[\{X^6\}\ \{X^7\}\right]^T \tag{2.29}$$

The zero order moment is always 1, since $\{X^0\} = \sum_{x=0}^{\infty} \frac{x!}{(x-0)!} P(x,t) = \sum_{x=0}^{\infty} P(x,t) = 1$.

Constant vector $\boldsymbol{\mu_c}$ and matrices A & A' are presented in equations (2.30), (2.31) and (2.32), respectively:

$$\boldsymbol{\mu_c} = [k_4\ 0\ 0\ 0\ 0]^T \tag{2.30}$$

$$A = \tag{2.31}$$
$$\begin{bmatrix} -k_3 & k_2 & -k_1 & 0 & 0 \\ 2k_4 & 4k_2 - 2k_3 & 2k_2 - 4k_1 & -2k_1 & 0 \\ 0 & 6k_2 + 3k_4 & 12k_2 - 6k_1 - 3k_3 & 3k_2 - 12k_1 & -3k_1 \\ 0 & 0 & 24k_2 + 4k_4 & 24k_2 - 24k_1 - 4k_3 & 4k_2 - 24k_1 \\ 0 & 0 & 0 & 60k_2 + 5k_4 & 40k_2 - 60k_1 - 5k_3 \end{bmatrix}$$

$$A' = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -4k_1 & 0 \\ 5k_2 - 40k_1 & -5k_1 \end{bmatrix} \tag{2.32}$$

In principle, any CME can be transformed into moment equations, through Z-transform. At this point instead of the CME there is a finite set of ordinary differential equations, albeit one that is not solvable because of the non-zero elements in A'.

2.3 ZI Closure Scheme

2.3.1 Solving Moment Equations by Maximizing the Information Entropy

ZI closure scheme can resolve the numerical closure issues of solving moment equations. The method is based on the maximum entropy principle. Since, Schlögl model is a single component network, the information entropy (equation (2.1)) is given by:

$$H = -\sum_{x=0}^{\infty} P(x) \ln P(x). \tag{2.33}$$

For factorial moment base (equation (2.15)), the Lagrangian (equation (2.2)) is given by:

$$\Lambda = H - \sum_{j=0}^{M} \lambda_j \left[\sum_{x=0}^{\infty} \frac{x!}{(x-j)!} P(x) - \{X^j\} \right], \tag{2.34}$$

Since the entropy is maximum, for every value $x = y$:

$$\frac{\partial \Lambda}{\partial P(y)} = 0 \tag{2.35}$$

or from equation (2.34):

$$\frac{\partial H}{\partial P(y)} - \frac{\partial}{\partial P(y)} \left\{ \sum_{j=0}^{M} \lambda_j \left[\sum_{x=0}^{\infty} \frac{x!}{(x-j)!} P(x) - \{X^j\} \right] \right\} = 0. \tag{2.36}$$

From equation (2.33):

$$\begin{aligned} \frac{\partial H}{\partial P(y)} &= \frac{\partial}{\partial P(y)} \left[-\sum_{x=0}^{\infty} P(x) \ln P(x) \right] \\ &= -\frac{\partial}{\partial P(y)} \left[P(y) \ln P(y) \right] = -\ln P(y) - 1. \end{aligned} \tag{2.37}$$

Additionally:

$$
\begin{aligned}
&\frac{\partial}{\partial P(y)}\left\{\sum_{j=0}^{M}\lambda_j\left[\sum_{x=0}^{\infty}\frac{x!}{(x-j)!}P(x)-\{X^j\}\right]\right\}\\
&=\sum_{j=0}^{M}\lambda_j\left\{\frac{\partial}{\partial P(y)}\left[\sum_{x=0}^{\infty}\frac{x!}{(x-j)!}P(x)\right]-\frac{\partial\{X^j\}}{\partial P(y)}\right\}\\
&=\sum_{j=0}^{M}\lambda_j\frac{y!}{(y-j)!}\frac{\partial P(y)}{\partial P(y)}=\sum_{j=0}^{M}\lambda_j\frac{y!}{(y-j)!}.
\end{aligned}
\tag{2.38}
$$

Combining equations (2.36), (2.37) and (2.38):

$$
P_H(y)=\exp\left[-1-\sum_{j=0}^{M}\lambda_j\frac{y!}{(y-j)!}\right] \tag{2.39}
$$

Equation (2.39) gives the stationary probability distribution for the maximum entropy solely in terms of the Lagrange multipliers. The subscript H emphasizes that the distribution is given for the maximum information entropy distribution.

Through the expression of the probability distribution for maximum entropy, both the lower-order and higher-order moments can now be related to a set of $M+1$ Lagrange multipliers. The following holds for every moment of the stationary probability distribution:

$$
\{X^m\}_H=\sum_{x=0}^{\infty}\frac{x!}{(x-m)!}P_H(x)=\sum_{x=0}^{\infty}\frac{x!}{(x-m)!}\exp\left[-1-\sum_{j=0}^{M}\lambda_j\frac{x!}{(x-j)!}\right] \tag{2.40}
$$

For example, the third moment of the Schlögl model example for up to $M=5$ lower-order moments is given by:

$$\{X^3\}_H = \sum_{x=0}^{\infty} \frac{x!}{(x-3)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right]$$
$$= \sum_{x=0}^{\infty} x\,(x-1)\,(x-2) \exp\left[-1 - \lambda_0 - \lambda_1 x - \lambda_2 x\,(x-1) - \lambda_3 x\,(x-1)\,(x-2)\right.$$
$$\left. -\lambda_4 x\,(x-1)\,(x-2)\,(x-3) - \lambda_5 x\,(x-1)\,(x-2)\,(x-3)\,(x-4)\right] \tag{2.41}$$

Based on equation (2.40), vectors $\boldsymbol{\mu}$ and $\boldsymbol{\mu'}$ of the Schlögl model for up to five lower-order moments can now be written only in terms of the six Lagrange multipliers:

$$\boldsymbol{\mu}_H = \begin{bmatrix} \sum_{x=0}^{\infty} \frac{x!}{(x-1)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right] \\ \sum_{x=0}^{\infty} \frac{x!}{(x-2)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right] \\ \sum_{x=0}^{\infty} \frac{x!}{(x-3)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right] \\ \sum_{x=0}^{\infty} \frac{x!}{(x-4)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right] \\ \sum_{x=0}^{\infty} \frac{x!}{(x-5)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right] \end{bmatrix} \tag{2.42}$$

$$\boldsymbol{\mu}'_H = \begin{bmatrix} \sum_{x=0}^{\infty} \frac{x!}{(x-6)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right] \\ \sum_{x=0}^{\infty} \frac{x!}{(x-7)!} \exp\left[-1 - \sum_{j=0}^{5} \lambda_j \frac{x!}{(x-j)!}\right] \end{bmatrix} \tag{2.43}$$

Matrices A and A' still have the same form (equations (2.31) and (2.32)). The difference is that now equation (2.6) depends only on a set of six Lagrange multipliers. Given that there is a set of five coupled equation, there is a need of one normalization constraint. This constraints comes from the fact that the zero$^{\text{th}}$ moment is constant and the zero$^{\text{th}}$ Lagrange multiplier depends on the rest of them. Since, the sum of probability distribution is 1 among all the state space, from equation (2.39):

$$
\begin{aligned}
\sum_x P_H(x) &= 1 \\
\sum_x \left[\exp\left(-1-\sum_{j=0}^{5}\lambda_j \frac{x!}{(x-j)!}\right)\right] &= 1 \\
\exp\left(-1-\lambda_0\right)\sum_x \left[\exp\left(-\sum_{j=1}^{5}\lambda_i \frac{x!}{(x-j)!}\right)\right] &= 1 \\
\lambda_0 = \log\left\{\sum_x \left[\exp\left(-\sum_{j=1}^{5}\lambda_j \frac{x!}{(x-j)!}\right)\right]\right\} &- 1
\end{aligned} \tag{2.44}
$$

Equation (2.44) consists of a major update compared to the algorithm described by Smadbeck & Kaznessis [28]. By recognizing that the zero$^{\text{th}}$ Lagrange multiplier is dependent on the rest of them, the numerical method becomes significantly more stable and always provides physically possible results. The system now has six equations, five for the moment equations and one from equation (2.44) and six unknown parameters, the Lagrange multipliers, and hence the probability distribution, can be computed using a root-finding numerical method such as Newton-Raphson. The algorithm of Newton-Raphson can be found in section 7.1 of Chapter 7.

2.3.2 Some Results for the Schlögl Model

Even though the Schlögl model is discussed in detail in Chapter 4, this section includes some characteristic results of the network using the updated ZI closure scheme algorithm.

For a specific set of kinetic constants, the stationary probability distribution calculated with ZI closure scheme, for a 12^{th} order closure ($M = 12$), is presented in figure 2.1. As the figure shows, ZI closure scheme accurately computes the steady state distribution compared to results from Gillespie's stochastic simulation algorithm (SSA) [25]. Table 2.1 also presents the first 8 moments calculated with ZI closure scheme and SSA, as well as the corresponding Lagrange multipliers calculated for ZI closure. The moment values calculated of the ZI closure scheme are close to the ones of SSA (forth column ot table 2.1).

Table 2.1: The second and third column of the table present the values of the moments of the stationary probability distribution of SSA and ZI closure scheme, respectively. The rows of the table indicate the order of the moment. The percentage difference between the moments calculated with the two methods relative to SSA results are located at the forth column. Fifth column presents the corresponding Lagrange multipliers calculated with ZI closure.

	SSA	**ZI closure**	**Difference (%)**	**Lagrange multiplier**
1st moment	40	40	6.0 10^{-6}	-1.1
2nd moment	2.4 10^3	2.4 10^3	4.5 10^{-6}	1.5 10^{-1}
3rd moment	1.6 10^5	1.6 10^5	1.0 10^{-5}	-8.5 10^{-3}
4th moment	1.1 10^7	1.1 10^7	1.2 10^{-5}	2.9 10^{-4}
5th moment	7.9 10^8	7.9 10^8	1.3 10^{-5}	-6.2 10^{-6}
6th moment	5.7 10^{10}	5.7 10^{10}	1.3 10^{-5}	9.1 10^{-8}
7th moment	4.2 10^{12}	4.2 10^{12}	1.3 10^{-5}	-8.9 10^{-10}

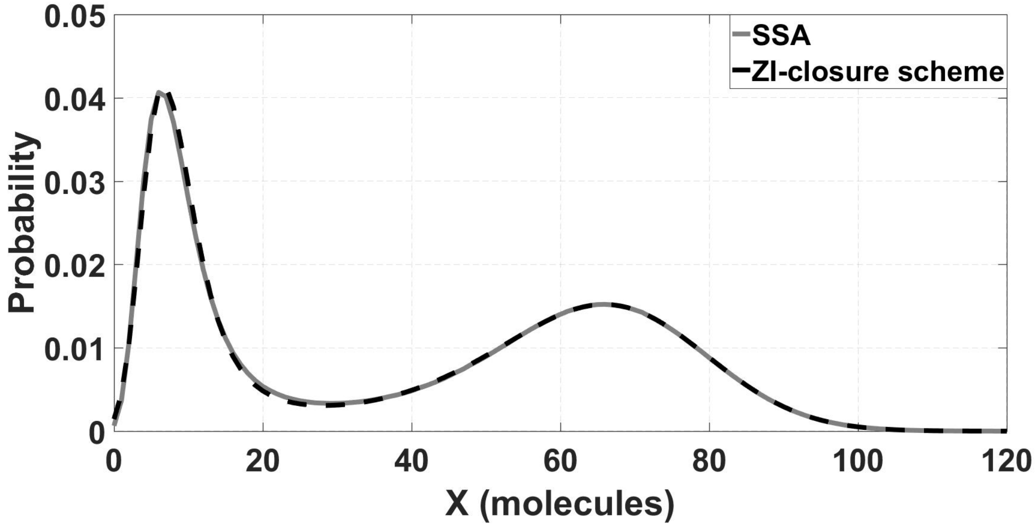


Figure 2.1: The figure compares the solution of ZI closure scheme for 12^{th} order closure and the solution of SSA for at least 500,000 trajectories. For this figure, table 2.1 and figure 2.2 the kinetic constants are $(k_1, k_2, k_3, k_4) = \Big(0.0015(molecules^2 \cdot s)^{-1}, 0.15(molecules \cdot s)^{-1}, 3.5s^{-1}, 20 molecules \cdot s^{-1}\Big)$.

Finally, figure 2.2 shows how the solution of ZI closure scheme is affected by the chosen order of the lower-order moments M. The number of lower-order moments affect the accuracy of the method and the figure demonstrated that at least 10 moments are needed to accurately reconstruct the probability distribution. After a specific value of lower-order moments ($M = 10$), the method converges to the same probability distribution.

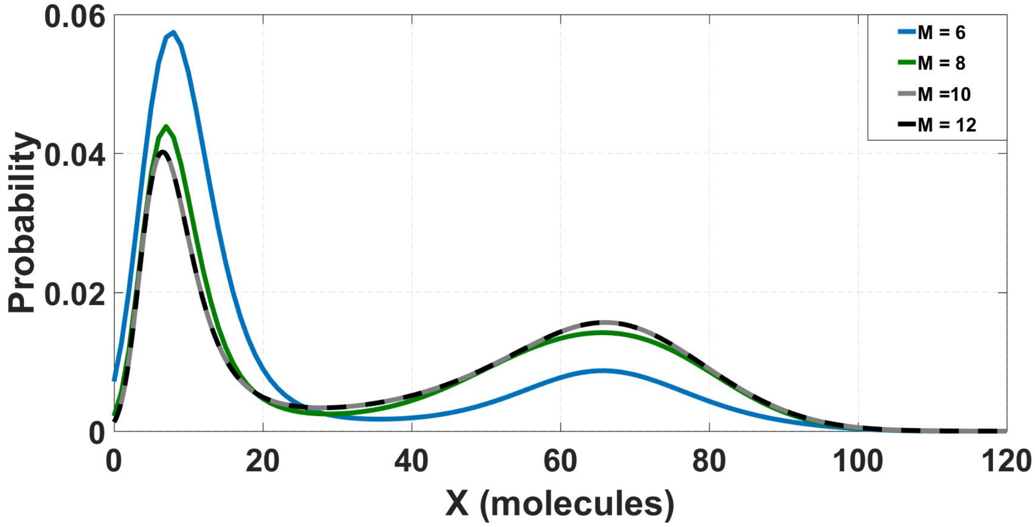


Figure 2.2: The stationary probability distribution of ZI closure scheme for different closure orders, M.

2.4 Conclusions

In this chapter, we explored in detail the use of factorial moments and the updated ZI closure scheme for solving moment equations. The Schlögl model is used as a simple, single-component system to demonstrate the mathematical formalisms and numerical results.

There are numerous challenges we face when using ZI closure. An important one is that there is currently no way to know *a priori* how many moments suffice for a numerically accurate reconstruction of the probability distribution. Furthermore, the reconstruction itself becomes a progressively more challenging task as the number of moments increases. ZI closure is certainly not an efficient method to compute the dynamics of probability distribution in non-steady state systems. This is because, at each time step an entropy maximization step must be implemented. Although Lagrange multiplier values from the previous time step can be used as an initial condition, the numerical implementation is slow. Additionally, ZI closure scheme assumes that each reaction network is an open system, which is a system with the minimum number of components, and each reaction has an elementary reaction rate.

An advantage of ZI closure over stochastic simulation algorithms is that it calculates steady state probability distributions directly, without resorting to simulations in time. Stochastic simulations must start with a specific initial condition in time for an ensemble of trajectories. Each trajectory can eventually reach a steady state. Instead, ZI closure scheme algorithm is initiated with a specific initial guess for the stationary probability distribution (*e.g.* a delta function) and numerically converges to the steady-state probability distribution. Furthermore, with ZI closure the tools become available to explore the stability of reaction network steady states and the sensitivity of the behavior to changes in system parameters [40,48], paving another avenue for investigating stochastic reaction networks.

Chapter 3

ZICS: a Windows and Matlab Application for ZI Closure Scheme

Zero-Information (ZI) closure scheme, as described in Chapter 2, is an innovative and promising approach in the field of stochastic kinetics. However, the algorithmic implementation of the method can be computationally challenging to a non-expert. In order to make the method more broadly accessible and foster the progress in the field, we have translated the ZI closure scheme into a user friendly application. The application, named ZICS, is a Windows standalone program. It accurately solves multidimensional multistable non-linear reaction networks. ZICS in publicly available at GitHub (https://github.com/mvlysidis/ZICS).

The purpose of this chapter is to introduce all the aspects of the application and educate the reader on how to solve any stochastic network with it.

3.1 Introduction

ZICS was created through Matlab's app designer. The application runs on Windows machines independent of Matlab installation. To ensure this, the ZICS installation package includes Matlab runtime. For non-Windows machines, the application runs through Matlab, version 2016b and later.

ZICS calculates the stationary probability moments of a reaction network and reconstructs the probability distribution. It is also able to generate the factorial moment equations. The application assumes that each reaction has an elementary reaction rate. ZICS can only be applied to open systems, which are networks with only the minimum number of independent components. Later in this chapter (section 3.3), we present an example how an non-open system can be transformed into one. The user has to specify the reaction network, kinetic constants and the number of lower-order moments (closure order).

For the construction of the reaction networks, the number of reactions and components are necessary. The definition of reaction networks has the form of matrices for reactant and product stoichiometries as well as the kinetic constants of each reaction. After the creation of the network, the user indicates the Newton-Raphson parameters. The limits of the state space for each component are specified. The maximum closure order is also required. The application can terminate the calculations when a new order closure does not improve the solution significantly, before the maximum order of closure is reached. Finally, the user can change the initial guess of the Lagrange multipliers. The default guess is a uniform distribution. The application runs locally and requires an insignificant amount of computational time to produce accurate results.

Results for Wilhelm's network [49] (table 3.1), a two-component multistable nonlinear network, are presented in figure 3.1. Each component performs a different bistable behavior, that the application is able to capture at steady state. The applications results are as accurate as Gillespie's SSA [9] ones. Graphs can be customized through the application. At the source code folder on GitHub, solutions for more than ten non-linear networks with ZICS are included.

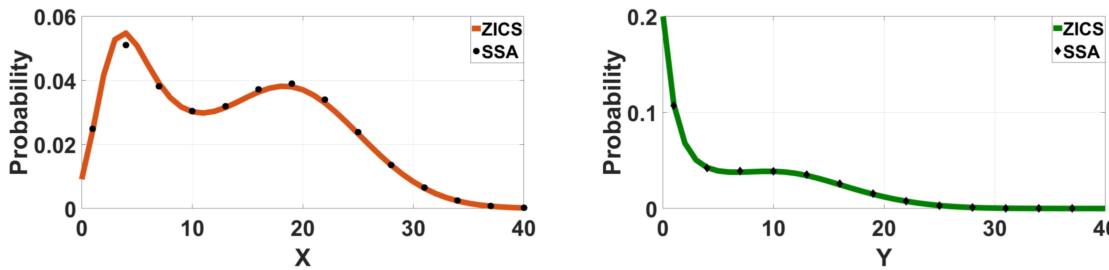


Figure 3.1: Stationary probability distribution of Wilhelm's network [49] as obtained with ZICS application. Left figure presents component X and right figure component Y. The results of compared with SSA ones (solid dots). The kinetic constants are reported in table 3.1.

3.2 Manual

This section presents the main features of ZICS application with the goal to educate the reader on how to use the application. In order to demonstrate the application's aspects and capabilities, Wilhelm's stochastic reaction network [49] is employed (table 3.1).

Table 3.1: The table shows the reactions for Wilhelm's bistable model and its kinetic constants. The network is used as an example to explain the application components.

Wilhelm's Reaction Network	**Kinetic constants**
$Y \xrightarrow{k_1} 2X$	$k_1 = 35$
$2X \xrightarrow{k_2} X + Y$	$k_2 = 1$
$X + Y \xrightarrow{k_3} Y$	$k_3 = 1$
$X \xrightarrow{k_4} \emptyset$	$k_4 = 9.74$
$\emptyset \xrightarrow{k_5} X$	$k_5 = 30$

3.2.1 Input the Network

In order to start solving the system, the user needs to input the reaction network and the kinetic constants into the application. The network is represented by the stoichiometric matrices for the reactants and products. Each reaction should be an irreversible reaction. Figure 3.2 presents the starting tab ("Input for Reaction Network") of the application.

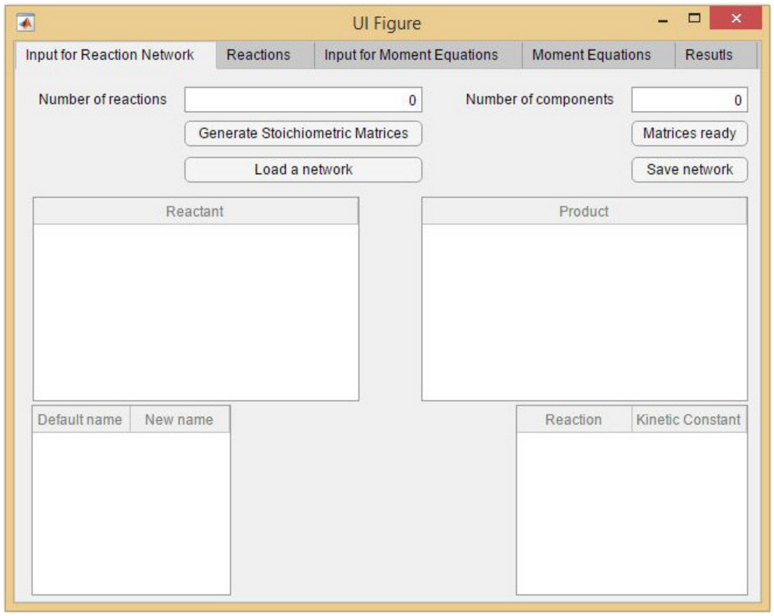
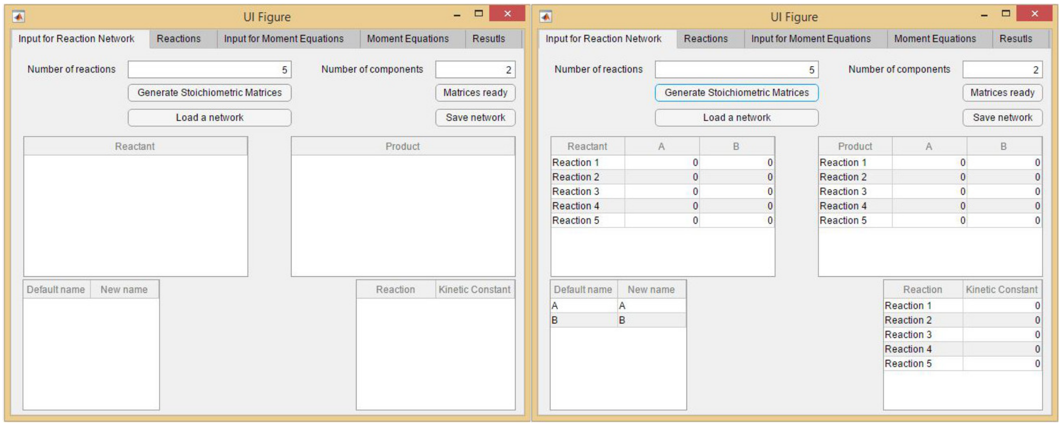


Figure 3.2: First tab of the ZICS app.

The number of reactions and components are the first options to be defined. Each reaction represents an irreversible reaction, thus a reversible reaction is represented as two separate ones. The number of components includes both the reactants and products. For the network at table 3.1, the number of reactions is 5 and components 2 (X and Y). The user can insert the number of reaction in the top left box and the number of components in the top right (figure 3.3a).



(a) (b)

Figure 3.3: (a) The number of reactions (5) denoted at top left box and the number of components (2) at the top right box. (b) The application creates blank matrices for the network after the ”Generate Stoichiometric Matrices” button is pressed.

After the number of components and reactions have been specialized, the user can create the reaction network. The user should press the ”Generate Stoichiometric Matrices” button and the application creates blank matrices for the reactants, products, name of components and kinetic constants (figure 3.3b).

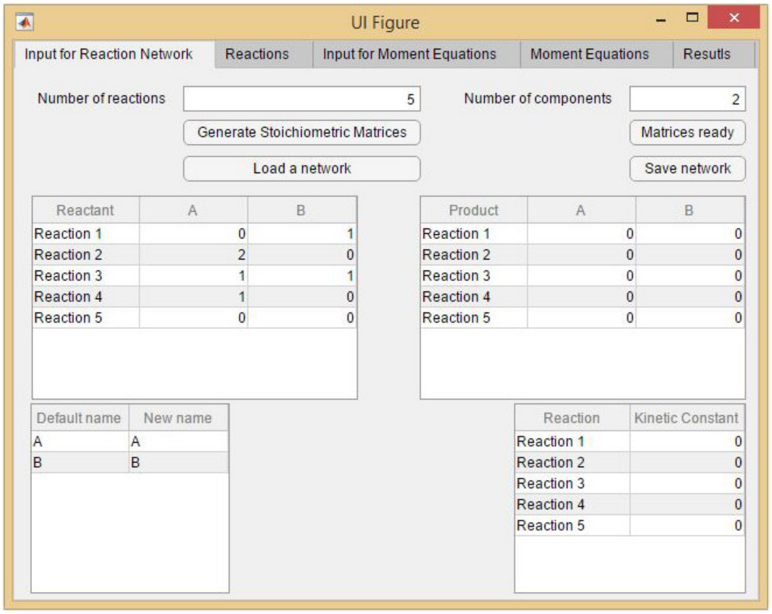


Figure 3.4: The reactants stoichiometric table (top left) is filled. Each row represents a reaction and each column a component. The table is filled according the reactions of table 3.1. Column A is used for component X and column B for component Y.

The user can now start inputing the reaction network. The user can first denote the stoichiometric matrix for the reactants in the top left table with the ”Reactant” indicator. The columns of the table denote the different components and the row the reaction of the network. In this case, we have two columns for the two components and five rows for the five reactions. We'll use column A for component X and column B for component Y. From table 3.1, it is easy to see that there are no molecules of X and 1 molecule of Y involved in the left side of the reaction. Thus, the first row-first column box has 0 and the first row-second column has 1. At the second reaction, 2 molecules of X and none of Y react. So the second row-first column box has 2. Following a similar procedure the user can fill the rest of the reactants table (figure 3.4).

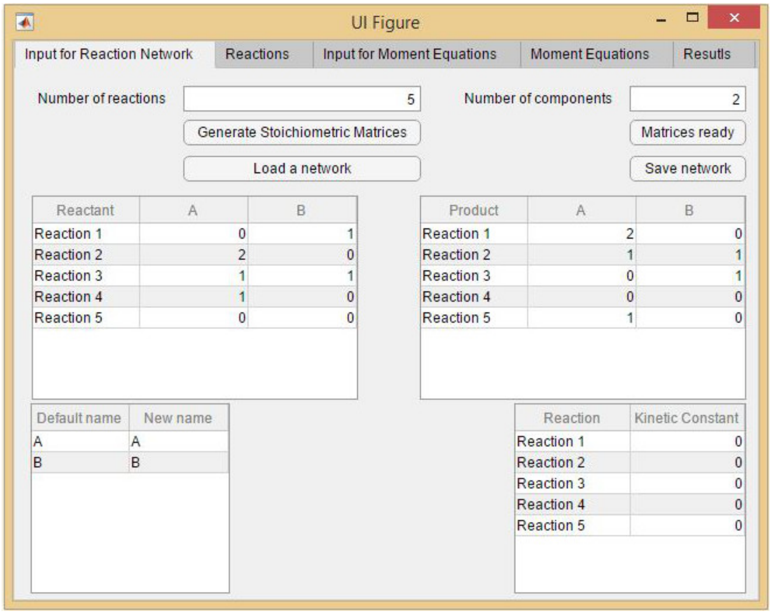


Figure 3.5: The products stoichiometric table (top right) is filled. Each row represents the same reaction and each column the same component as in the reactants stoichiometric matrix. The table is filled according the reactions of table 3.1.

Similarly, the user can input the stoichiometric matrix for the products in the top right table with the "Product" indicator. Each row of the table represents a reaction. The order of the reactions does not affect the solution, however each row of the reactants and products matrices should represent the same reaction. The columns of the table represent the components. In this case, column A refers to component X and column B to component Y. Again, the pairing of columns and components does not influence the final result, however each column of the reactants and products matrices should represent the same component. For the reaction network presented at table 3.1, the first reaction produces two molecules of X and zero molecules of Y. Thus, the first row-first column box is filled with 2 and the first row-second column with 0. The second reaction produces one molecule of X and Y, hence the second row-first column box is filled with 1 (for X) and the second row-second column box has also 1 (for Y). With the same strategy, the rest of the products stoichiometric matrix can be filled (figure 3.5). At this point the reaction network has been fully defined and the user can input the reaction kinetic constants.

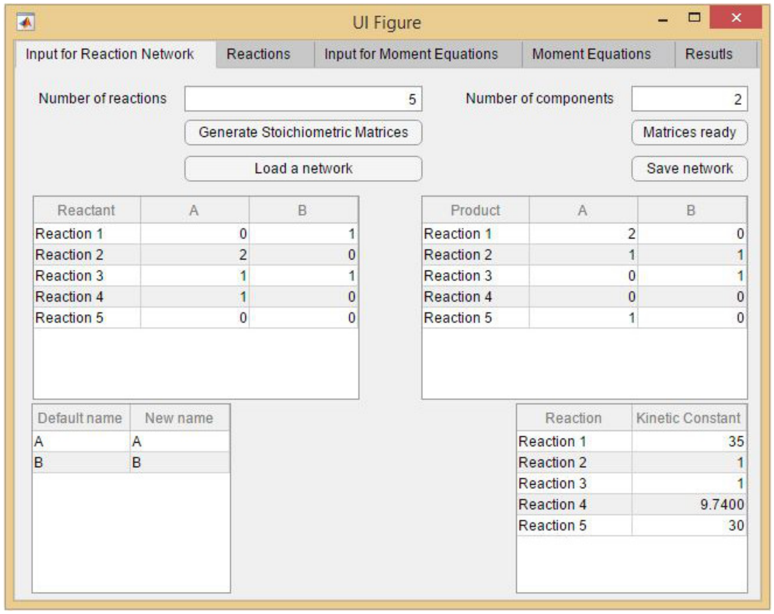


Figure 3.6: The kinetic constants table (bottom right) has been filled. Each row represents the same reaction as in the stoichiometric tables for the reactants and products. The value of the kinetic constants is based on table 3.1.

The reaction constants can be inputted in the bottom right table indicated with ”Kinetic Constant”. The table has one row representing the value of the kinetic constants and multiple rows, one for each reaction. Each row should represent the same reaction as in the stoichiometric matrices for products and reactants. Based on table 3.1, the first element of the table should 35 (for the kinetic constant of the first reaction), the second 1, the third 1 etc. (figure 3.6).

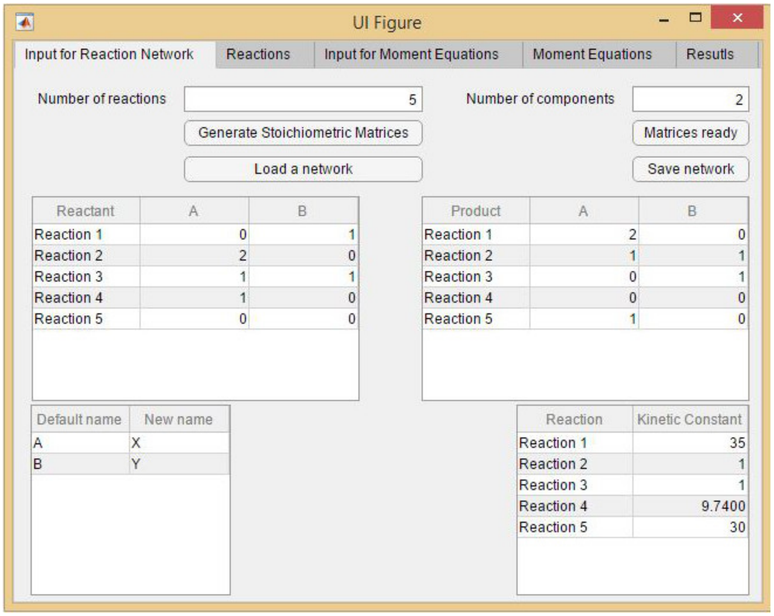


Figure 3.7: At the bottom left table, the user can rename the components. A and B represent X and Y, respectively.

Finally, the application gives the ability to the user to rename the reactions components. The default option is the letters of the alphabet. The user can rename the components at the bottom left table. In this example, component A represents X so it is renamed as X and component B as Y, as shown in figure 3.7.

The user has also the ability to save the network. The ”Save network” button saves the network in a Matlab .mat format. The application can load the previously saved networks with the ”Load a network” button. This function allows to load sbml files (.xml) in a .tsv format. In order to first transform the .xml files to .tsv, the user is directed to free available website: `https://rumo.biologie.hu-berlin.de/SBtab/default/converter`.

After the network, the kinetic constants and the name of the reactants have been inputted, the user can move on by pressing ”Matrices Ready” button. The user has the ability to check the form of the network at the second tab (”Reactions”), as shown figure 3.8. If there is an error in the network, the user can go back to the previous tab (”Input for

Reaction Network") and make the appropriate changes as described above. If changes in the "Input for Reaction Network" tab has been made, the user should press the "Matrices Ready" button to make them final and apply them for the rest of the application.

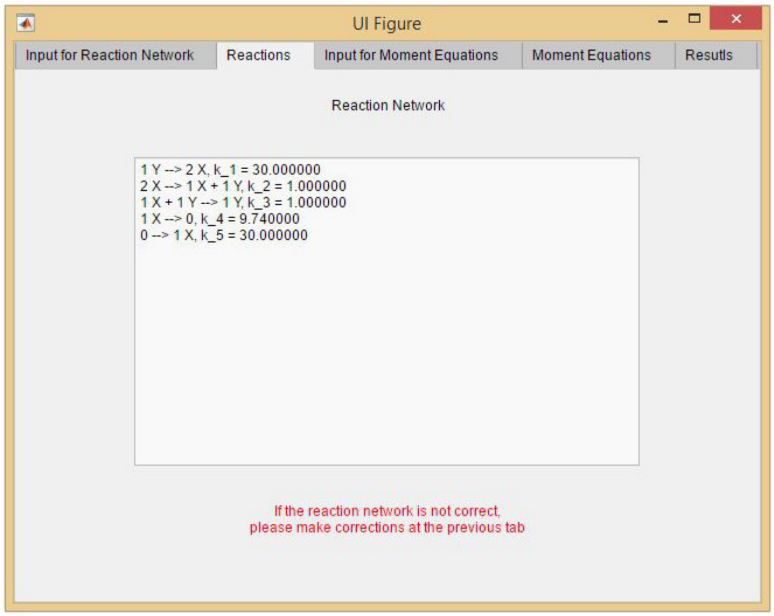


Figure 3.8: At the second tab named "Reactions" the user can check the form of the network. In case, an error has occurred the user can go back to the previous tab "Input for Reaction Network" and make the appropriate changes.

3.2.2 Simulation Parameters

After the input of the reaction network and pressing the "Matrices ready" button, the user can enter the simulation parameters in order to solve the system. Simulation parameters can be entered at the third tab (named "Input for Moment Equations" as shown in figure 3.9). More details about the code of the application can be found in the section 7.1 of Chapter 7.

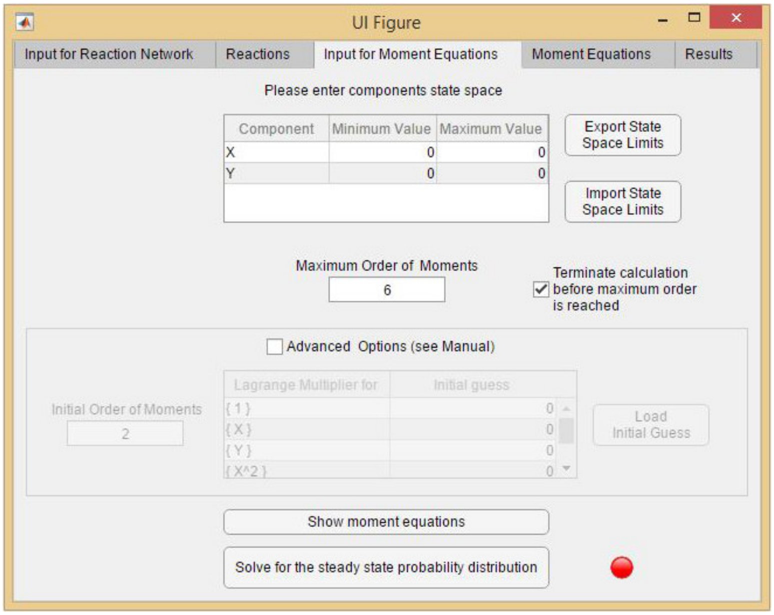


Figure 3.9: The simulation parameters can be entered at the ”Input for Moment Equations” tab.

The user can first specify the state space of the system. The state space of the system indicates the minimum and maximum number of molecules for each component. ZICS application solves stochastic reaction networks numerically and thus the state space of system should be indicated by the user and cannot be infinite. In the application, the state space is represented as a table (figure 3.9). Each row represents each of the components. There are two columns, one for the minimum value of the state space and one for the maximum value.

It is suggested that the minimum value of each component is 0. The maximum value can vary per system and component. If the maximum value is less than the actual value, the application will produce inaccurate results, since it was not allowed to perform calculation in the whole necessary space. It is suggested that the maximum value is slightly higher than the actual value so that the application can perform calculations at the whole state space. There are two common approaches in case the user does not have a good educated guess about the maximum value. The first is to input an arbitrarily high numerical value. This can ensure accurate results, however the computational time required to produce results increases with the values of the state space. Thus, this approach can be computationally costly. An alternative way is to solve the system with a relatively average numerical value and then solve the the system with a higher maximum value and compare the two solutions. The final correct solution of the network should be independent of the state space values. For this example, we chose to use 50 molecules as the maximum value for component X and 40 for component Y as shown in figure 3.10. The user has the ability to save the state in a Matlab .mat format by pressing the ”Export State Space Limits” button and load them again by pressing the ”Import State Space Limits” button.

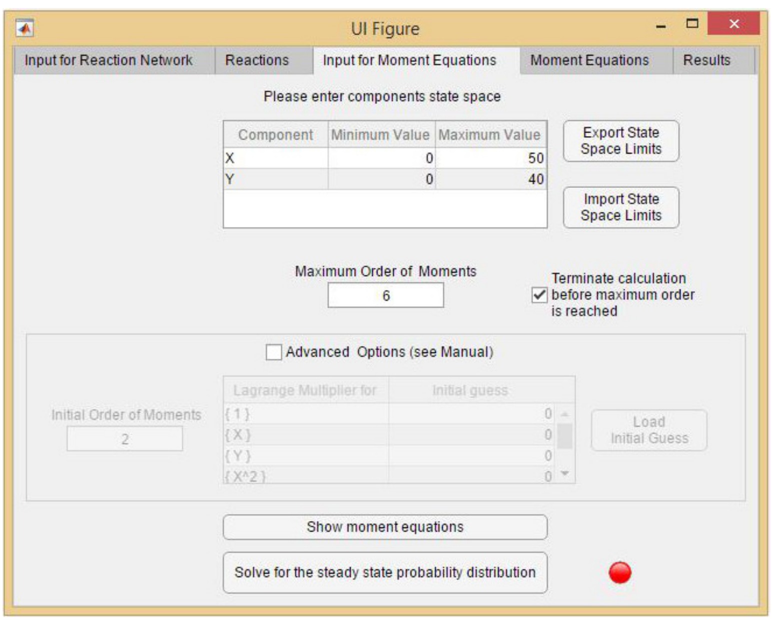


Figure 3.10: The state space for component X is between 0 and 50 and for component Y is from 0 to 40.

After the state space is specified, the user can enter the maximum order of moments. The application solves moment equations and the number of moments can change the size of the equations solved and thus the accuracy of the method. The maximum order of moments can be entered at the homonymous box (figure 3.9). For the majority of networks tested, an order of 6 moments is enough. The application has the ability to dynamically change the number of moments and test the accuracy of the results. Thus, the application can terminate the calculations before the maximum number of moments is reached if the solution meets the desired accuracy. This mode is automatically enabled. If desired, the user can disable this feature and force the application to perform calculations until the maximum number of moments is reached, by unselecting the box indicated as ”Terminate calculation before maximum order is reached”. It is suggested that the user inputs a high number of maximum moments and keeps the feature enabled. For this example, we chose that 8 order of moments is large enough (figure 3.11).

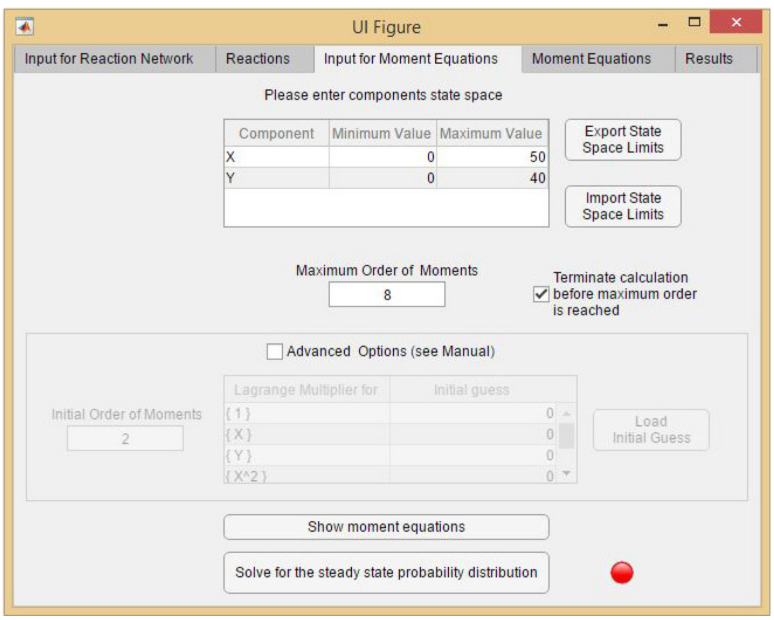


Figure 3.11: The maximum order of moments can be modified. The application has the ability to terminate calculations before this number is reached. In this case, the maximum order of moments was changed to 8.

With the order of moments and state space specified, the application is ready to produce results. The application calculates the stationary probability distribution for each component and their stationary probability moments. In order to start the calculations, the user should press the ”Solve for the steady state probability distribution” button. Every time changes are performed in the ”Input for Moment Equations” tab, the ”Solve for the steady state probability distribution” button should be pressed to apply those changes. There is a lamp next to this button that indicates the status of the calculations. Before or during the calculations, the light of the lamp is red (figure 3.11). The light becomes green when the program has finished running and the stationary probability distribution for all the components of the network have been calculated (figure 3.12).

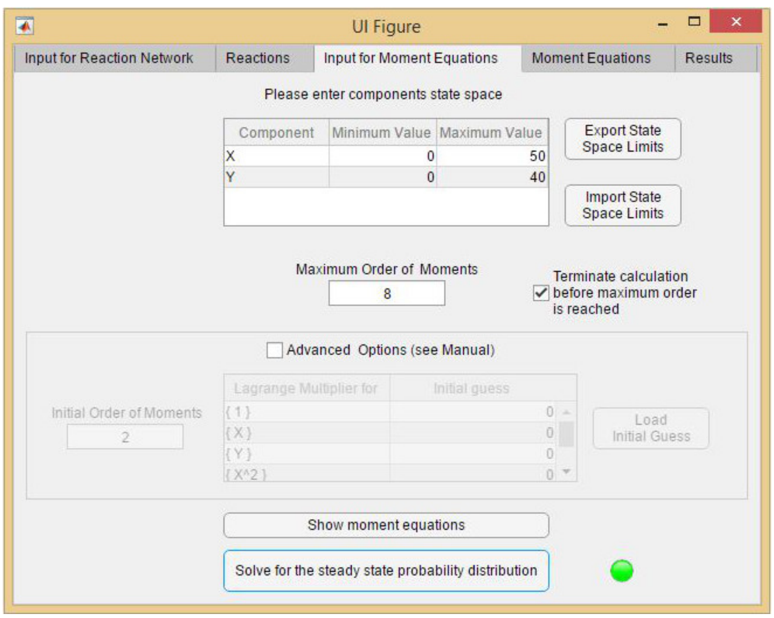


Figure 3.12: Calculations start when the ”Solve for the steady state probability distribution” button is pressed. When the lamp next to the button is green, the calculations have finished.

Advanced Options

ZICS application is using ZI-closure scheme’s Newton-Raphson algorithm. As such, the program requires an initial guess; the initial guess comes in the from of the initial number of moments and their associated Lagrange multipliers. For more information about how the Lagrange multipliers affect the algorithm of the application, the reader is directed to the section 7.1 of Chapter 7 or the book chapter [39]. The default initial condition is a uniform distribution with second order of moments and zero for all Lagrange multiplier values. The user has the ability to change the initial guess of the program by checking the ”Advanced options (see Manual)” box. In the advanced options section, the user can modify the initial order of moments, the initial guess for Lagrange multipliers and also load an initial guess (by pressing the ”Load Initial Guess” button). It is encouraged this section to mainly be used by experienced users.

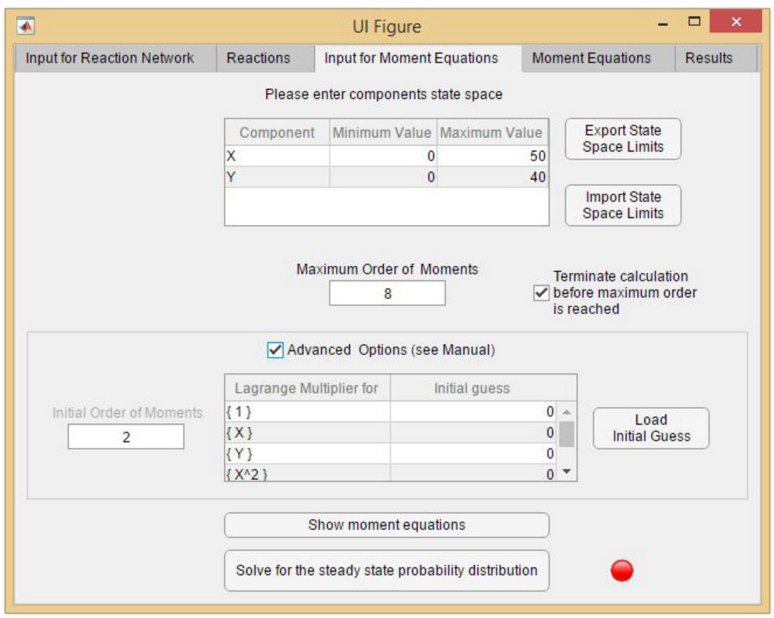


Figure 3.13: The user can input an initial guess for the algorithm by checking the ”Advanced options (see Manual)” box.

3.2.3 Results

Stationary Probability Distribution

The results of the application are displayed at the last tab named ”Results” (figure 3.14). The stationary probability distribution for the first component is automatically plotted at the top center. The x-axis represents the number of molecules of the component and the y-axis the probability. The plot can be exported by pressing the ”Save plot” Button in multiple formats (.fig, .jpg, .png, .eps). Below the plot, the stationary values of the moments as well as their Lagrange multipliers are presented. The results can be saved in a Matlab (.mat) or spreadsheet (.xlsx) format by clicking the ”Export Results” button. There is also the option to save the stationary Lagrange multipliers in a Matlab (.mat) format. Their values can be used at the advanced options section (section 3.2.2) for future runs.

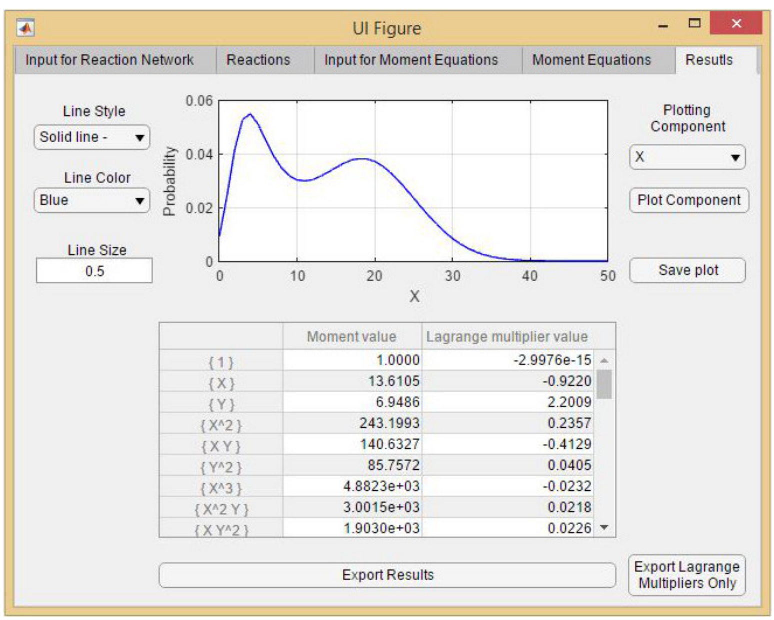


Figure 3.14: At the last tab named ”Results” the user can find the solution of the network. The stationary probability distribution of the first component is automatically plotted. Additionally, the stationary moments and Lagrange multipliers are reported.

The user has the ability to plot the stationary probability distribution for all the components of the system. By clicking the ”Plotting Component” drop-down menu, the user can select the desired component. In order to plot the new component, the ”Plot Component” button should be pressed (figure 3.15). The user can also choose between four different line styles, seven colors and change the line size of the plot from the respective options at the left-hand side of the plot (figure 3.16). Again, to apply any changes the user should press the ”Plot Component” button. Any plots can be saved in four different formats (.fig, .jpg, .png or .eps) by pressing the ”Save plot” Button.

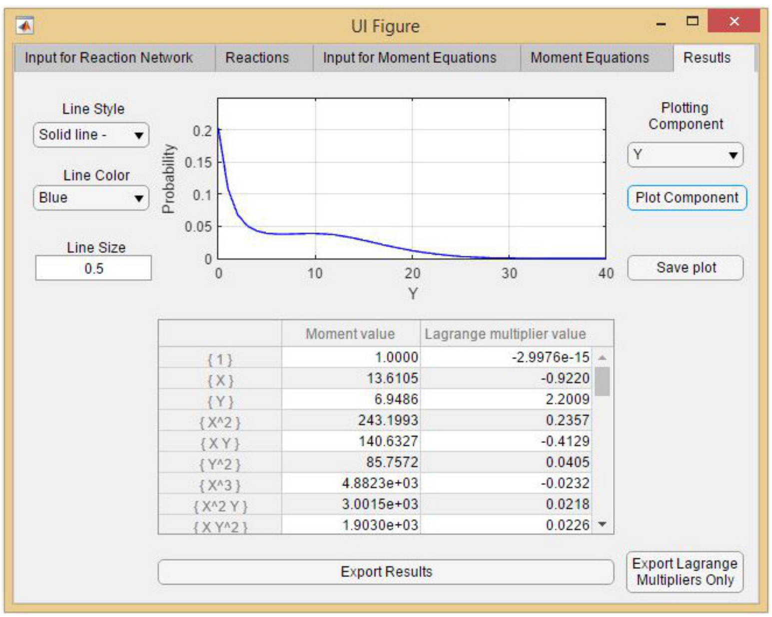


Figure 3.15: The ”Plotting Component” drop-down menu gives the ability to plot different components. Here, the stationary probability distribution for component Y is plotted after the press of ”Plot Component” button.

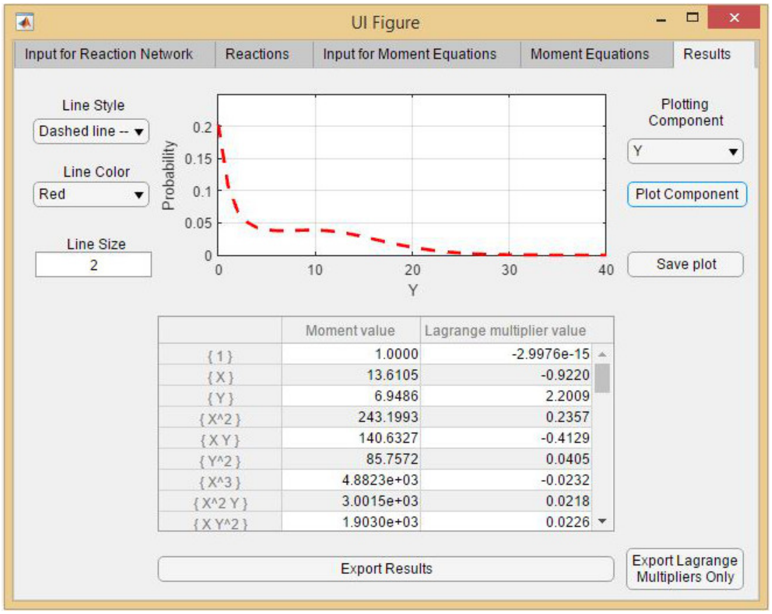


Figure 3.16: The user can modify the plot with the options of the left upper side and press the ”Plot Component” button. Here, the ”Line Style” is chosen to be dashed, the ”Line Color” red and the ”Line Size” 2.

Moment Equations

Aside from calculating the stationary probability distribution of stochastic networks, the application also calculates the moment equations and its associated matrices. The moment equations are displayed at the ”Moment Equations” tab, as shown in figure 3.17. In order for the matrices to appear, the ”Show moment equations” button should be pressed at the ”Input Moment Equations” tab (figure 3.12). This function is independent of the function that calculates the stationary probability distributions. The user can calculate the moment equation without solving the system. The moment equations can be exported in a Matlab (.mat) or spreadsheet (.xlsx) format by clicking the ”Export Matrices” button.

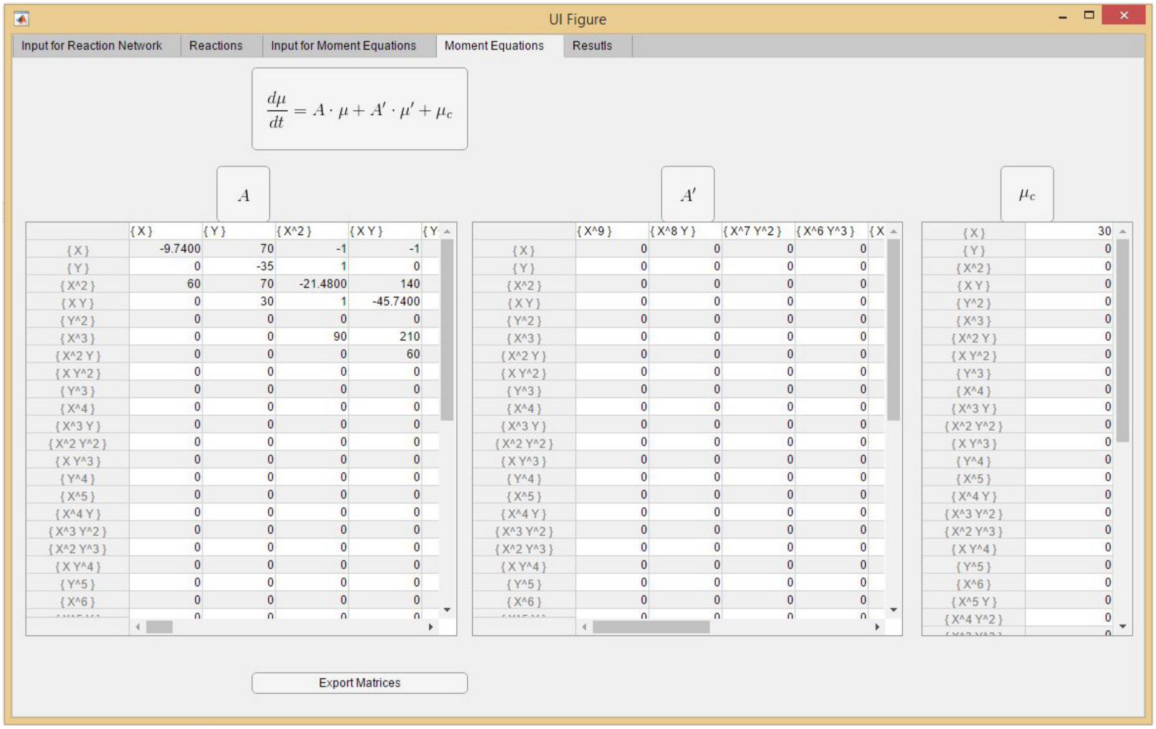


Figure 3.17: Moment equations matrices are displayed at the "Moment Equations" tab. This function is possible by pressing the "Show moment equations" button at the "Input Moment Equations" tab and independent of the function that calculated the stationary probability distribution.

3.3 The Michaelis-Menten Reaction Network

ZICS application is able to only solve open systems, which are systems with only the minimum possible number of components. Any closed system needs to be transformed into an open one in order to be solved with the application. In this section, the Michaelis-Menten reaction network [50, 51] is used as an example of transforming a closed system to an open one. The closed form of Michaelis-Menten can be found in table 3.2.

The network has 2 degrees of freedom (components S and E), however they are 4 components present (S, E, $S:E$ and P). The reaction network needs to be transformed into an open system based on the two relations: $E_{total} = E_T = S:E + E$ and $S_{total} =$

Table 3.2: Michaelis-Menten network. Left columns show the networks closed form and the right ones show the open system.

Closed System		Open System	
Reactions	*Kinetic Constants*	*Reactions*	*Kinetic Constants*
		$S + E \longrightarrow \emptyset$	k_1
$S + E \longrightarrow S:E$	k_1	$\emptyset \longrightarrow S + E$	$E_T\ k_2$
$S:E \longrightarrow S + E$	k_2	$E \longrightarrow S + 2E$	$-k_2$
$S:E \longrightarrow P + E$	k_3	$\emptyset \longrightarrow E$	$E_T\ k_3$
		$E \longrightarrow 2E$	$-k_3$

$S_T = S : E + S + P$. Thus, components $S : E$ and P depend on components S and E through $S : E = E_T - E$ & $P = S_T - S - S : E = S_T - S - E_T + E$. Each reaction that includes either component $S : E$ or P has to be reconstructed since these components are now dependent, and only independent components should remain in the reaction network.

If the dependent components are products of a reaction, they can simply be removed from the reaction with no loss of information. For example, reaction 1 of the closed system of table 3.2 (left column) is transformed from $S + E \xrightarrow{k_1} S : E$ to $S + E \xrightarrow{k_1} \emptyset$ with the same kinetic constant.

If the dependent components are reactants of a reaction, then the transformation of the reaction is as follows:

Reaction 2 of left column of table 3.2 is:

$$S : E \xrightarrow{k_2} S + E$$

The reaction can be transformed into:

$$\emptyset \xrightarrow{k_2 \cdot S:E} S + E$$

with a new kinetic constant of $k_2 \cdot S : E$. $S : E$ can be now substituted with $S : E =$

$E_T - E$ from the mass balances and hence the reaction is:

$$\emptyset \xrightarrow{k_2 \cdot (E_T - E)} S + E$$

with a new kinetic constant of $k_2 \cdot (E_T - E) = k_2 \cdot E_T - k_2 \cdot E$. Since, the kinetic constant has two parts ($k_2 \cdot E_T$ and $-k_2 \cdot E$), the reaction can be split into two reactions:

$$\emptyset \xrightarrow{k_2 \cdot E_T} S + E \qquad \emptyset \xrightarrow{-k_2 \cdot E} S + E$$

Now the reaction includes only independent components, however the kinetic constant does not have a constant value any more with the presence of component E. The user needs to remove component E from the kinetic constant and input it in the reaction as both a reactant and a product:

$$\emptyset \xrightarrow{k_2 \cdot E_T} S + E \qquad E \xrightarrow{-k_2} S + 2E$$

Now, closed reaction 2 has been transformed into two open reactions. With the same methodology, the reconstructed open Michaelis-Menten reaction network is displayed on table 3.2.

3.4 Conclusions

In this chapter, we presented ZICS, a standalone Windows application, for the calculation of the stationary probability distribution of stochastic biochemical reaction networks. The application is based on the Newton-Raphson algorithm of ZI closure scheme (Chapters 2 & 7, [39]). With minimal computational cost, it accurately solves multistable multidimensional networks. ZICS also offers the ability to generate the factorial moment equations of a given network. ZICS is designed with the Matlab App deisgner. For non-Windows machines, the application can be used through Matlab, version 2016b and later.

The application is open source and publicly available (github.com/mvlysidis/ZICS). At the source code folder, there are more than ten reaction networks that the application accurately solves. The networks vary in complexity, probability distribution modes and number of components, including the bistable Schlögl model [46], the multistable Wilhelm's system [49] and a five-component limit cycle. Most importantly, the application runs locally and requires an insignificant amount of computational time to produce accurate results.

Chapter 4

On Differences Between Deterministic and Stochastic Schlögl Model

4.1 Background

Chemical reaction kinetics have been canonically modeled with ordinary differential equations since the pronouncement of the law of mass action kinetics, 150 years ago [52]. This macroscopic, continuous-deterministic modeling formalism is appropriate at the thermodynamic limit, when the volume of the system and the numbers of molecules of reactants all tend to very large values.

Markov chain models can be used for chemical reactions away from the thermodynamic limit [8]. Models are then formulated in terms of discrete numbers of molecules for each of the chemical species present at time t. The system evolves stochastically, and the all-encompassing chemical master equation (CME) can model the probability distribution of the system being at a particular state at time t [5].

Kurtz [53, 54] explored the relationship between stochastic and deterministic models when the macroscopic equations have a unique, asymptotically stable solution, and demonstrated that the deterministic model is the thermodynamic limit of the stochastic one.

However, when the ordinary differential equations admit more than one stable solutions, the two formalisms may give starkly different results for even simple, small chemical reaction models [55–57]. The Schlögl model is such a simple model [7], described as follows [46]:

$$3\,X \underset{k_2'}{\overset{k_1}{\rightleftharpoons}} 2\,X + A \qquad\qquad X \underset{k_4'}{\overset{k_3}{\rightleftharpoons}} B.$$

This is a one-dimensional model. The only variable is the number of molecules X. Species A and B are present in the system with constant concentrations, arriving from separate external reservoirs. Setting $k_2 = k_2'\,[A]$ and $k_4 = k_4'\,[B]$, the model is simplified as

$$3\,X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2\,X \qquad\qquad X \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \emptyset.$$

For a large system of volume V, the concentration $x = X/V$ changes in time according to the law of mass action:

$$\frac{dx}{dt} = -k_1\,x^3 + k_2\,x^2 - k_3\,x + k_4 \tag{4.1}$$

This cubic equation admits either a single stable solution, or two stable solutions and one unstable one [58]. The solution depends on kinetic constant and reservoir concentration values. For the rest of the chapter, the values of all kinetic parameters are constant, except k_4. The values of the constant kinetic parameters are $(k_1,\ k_2,\ k_3) = (0.42\ (molecules^2 \cdot s)^{-1},\ 70\ (molecules \cdot s)^{-1},\ 3150\ s^{-1})$. We present results for k_4 ranging from $10^4\ molecules \cdot s^{-1}$ to 7 $10^4\ molecules \cdot s^{-1}$.

The chemical master equation for the Schlögl model has been derived in Chapter 2 and from [59]:

$$\frac{\partial P(X;t)}{\partial t} = \left[\frac{k_2}{V}(X-1)(X-2) + k_4 V\right] P(X-1;t)$$
$$-\left[\frac{k_2}{V}X(X-1) + k_4 V + \frac{k_1}{V^2}X(X-1)(X-2) + k_3 X\right] P(X;t)$$
$$+\left[\frac{k_1}{V^2}(X+1)X(X-1) + k_3(X+1)\right] P(X+1;t) \tag{4.2}$$

As discussed previously [28], for small system sizes the stationary probability distribution $P(X;t)$ is either unimodal or bimodal, depending on parameter values. Difficulties in computing the stationary probability distribution for mesoscopic systems, which are still under the influence of molecular fluctuations, but have molecular populations too large for stochastic simulation algorithms, have hampered the analysis of stationary probability distributions.

Herein the master equation is solved with the ZI closure scheme (Chapters 2, 3 & 7) for large system sizes that were previously unattainable with other methods. We keep the concentration x constant, while systematically increasing the system size, in order to investigate the collapse of bimodality for mesoscopic and large systems, and determine the limits of correspondence between CME and ODE models.

Note that the systems studied in this chapter are not at equilibrium. The chemical potentials of the two reservoirs in the Schlögl reaction are set to be different. There is then mass flowing from the reservoir with high chemical potential to the reservoir with low chemical potential. The flow of mass is at steady state, and as a result the probability distribution of X within the system is stationary. By varying the kinetic constants, the steady state changes. The focus in on the stationary probability distribution of these non-equilibrium steady states (NESS).

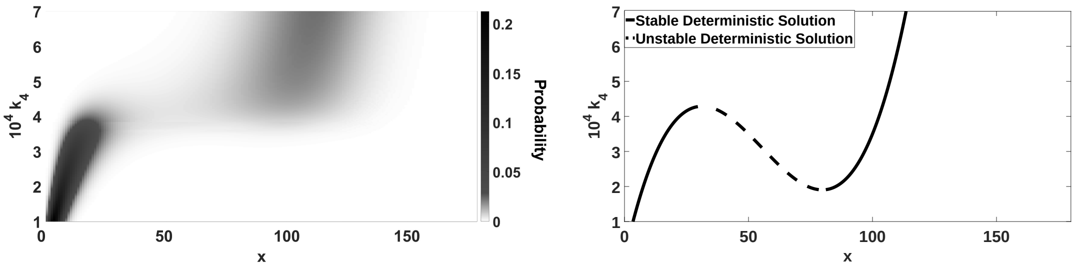
We study the Schlögl model as a simple example of a bistable system. There is a plethora of natural systems exhibiting bistability, such as the lysis and lysogeny system in phage [60]. Over the past two decades, there has been particular interest in both synthetic and natural biological systems exhibiting either temporal or spatial bistability [61–65]. The Schlögl model has been recognized as one of the simplest models that can capture essential elements of bistable behavior [49]. In principle, insights gained from studying this model can be applied to other bistable stochastic networks. Of particular concern when studying biological systems is the choice of an appropriate modeling formalism. This study may serve as a guide on the limitations of either stochastic or deterministic models used to study bistable systems in biology.

4.2 Results

This section presents results obtained with ZI closure scheme for the stochastic Schlögl model.

4.2.1 Bimodality Collapse for Mesoscopic Systems

In figure 4.1a, steady state probability distributions, $P(x)$, are plotted for a range of values of k_4.



(a) (b)

Figure 4.1: The figure shows the effect of kinetic constant k_4 on the stationary probability distribution for small systems with $V = 1$. Figure 4.1a shows the behavior of the stochastic system. Darker areas present higher probability. Figure 4.1b shows the behavior of the deterministic system. Solid lines represent the stable deterministic solutions and the dashed-dotted line represents the unstable solution of the ODE model. The other kinetic constants are: $(k_1,\ k_2,\ k_3) = (0.42,\ 70,\ 3150)$

In accord with the conclusion drawn by Kurtz, when the probability distribution is unimodal, the average of X corresponds to the single stable ODE solution. This is shown in the figure 4.1b, where the ODE solution has been plotted on the x-k_4 plane. When the distribution is bimodal, the peaks correspond to the two stable deterministic attractors (figure 4.1). As the kinetic constant increases, the system transitions from a unimodal distribution to a bimodal and then back to a unimodal one. A main purpose of this chapter is to explore deviations from this congruency, as the system size varies.

In the rest of the document, the terms 'probability peak (peak)' and 'attractor' will be used interchangeably. When the intention is to draw attention to the probability distribution and focus on the stochastic behavior the term 'peak' is preferable. On the other hand, when the intention is to compare the results of the deterministic and stochastic models, the term 'attractor', which pertains to both modeling formalisms, is preferred.

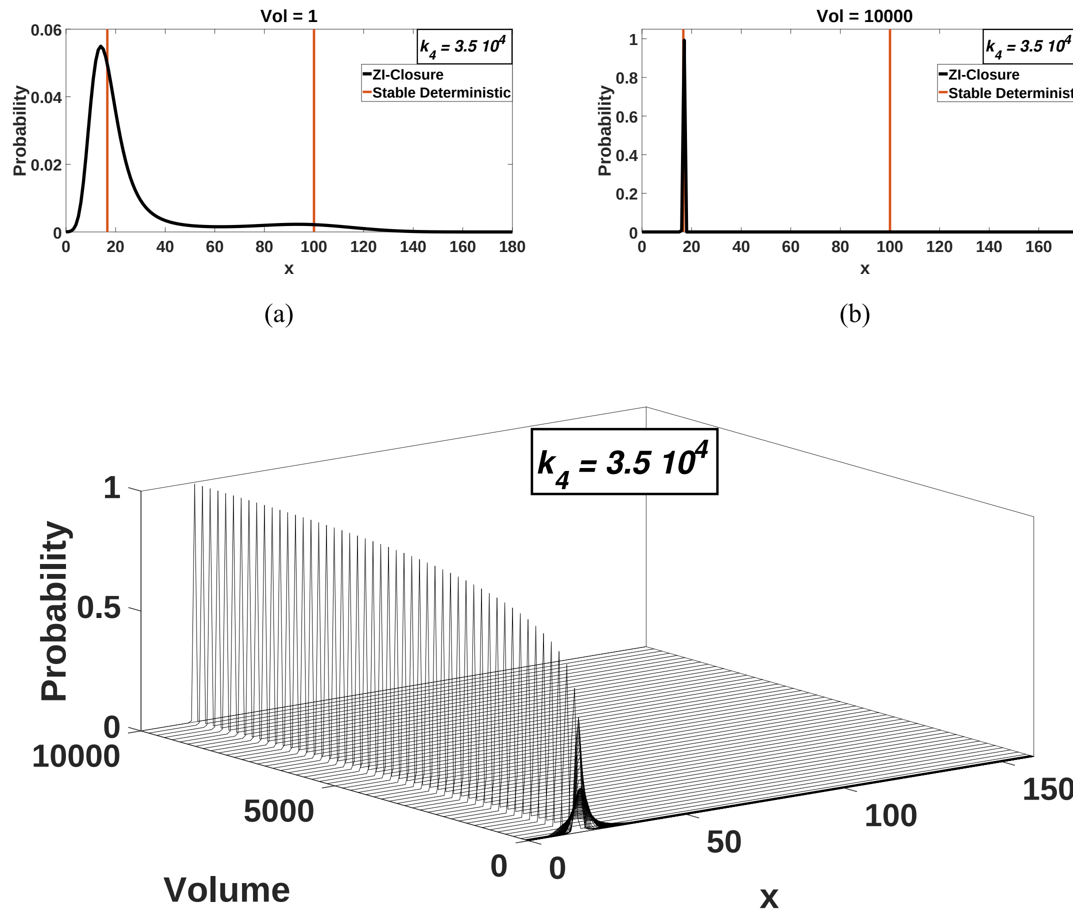
Parenthetically, it is interesting to note that the range of k_4 values where the stochastic model exhibits bimodality does not precisely correspond to the range of k_4 values where the deterministic model exhibits bistability. The reason for this minor discrepancy is that in the ODE model (equation (4.1)) higher-order reactions are represented in terms of X^2 and X^3, whereas in the CME (equation (4.2)), there are $X(X-1)$, $(X-1)(X-2)$, $(X+1)X(X-1)$ and $X(X-1)(X-2)$ terms.

When the system is bistable, only one of the solutions will be reached in finite time in the ODE model. Which one of the solutions will be reached depends on the initial conditions. In contrast, the stochastic model is ergodic and explores the entire state space (all the possible numbers of molecules of X) with a frequency proportional to the probability, visiting both attractors, regardless of initial conditions.

In other words, the stochastic model identifies and distinguishes all solution attractors, whereas the ODE model cannot. This is in accord with, among others, the study of Ge and Qian [66], who concluded that although deterministic differential equations can define numerous attractors, they provide no information on the relative probabilities between them. Only stochastic model solutions can provide such insight.

What is intriguing is that as the system size increases, the stochastic Schlögl model behavior ceases to correspond to the ODE one. The deterministic model solution is not dependent on the size of the system and will always exhibit the same bistable behavior, for certain ranges of parameter values. In the stochastic model, the size of the system plays a critical role. As the size increases, the bimodality can be destroyed and the stationary probability distribution can become a delta-like function [7, 55, 56, 67, 68].

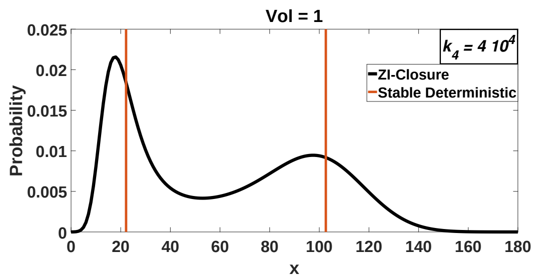
This behavior is observed in figures 4.2 & 4.3, where the stationary probability distribution is shown as a function of the volume for two values of k_4. Both k_4 values are within the range that produces two stable solutions for equation (4.1). For small volumes, the stochastic model stationary probability distribution exhibits bimodality. The peaks then approximately correspond to the ODE solutions (figures 4.2a & 4.3a).



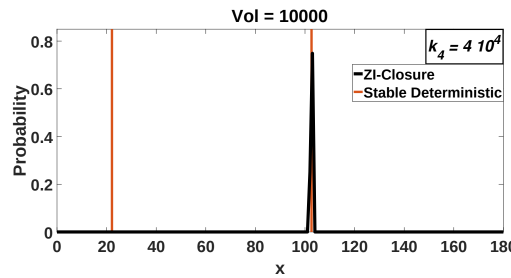
(a) (b)

(c)

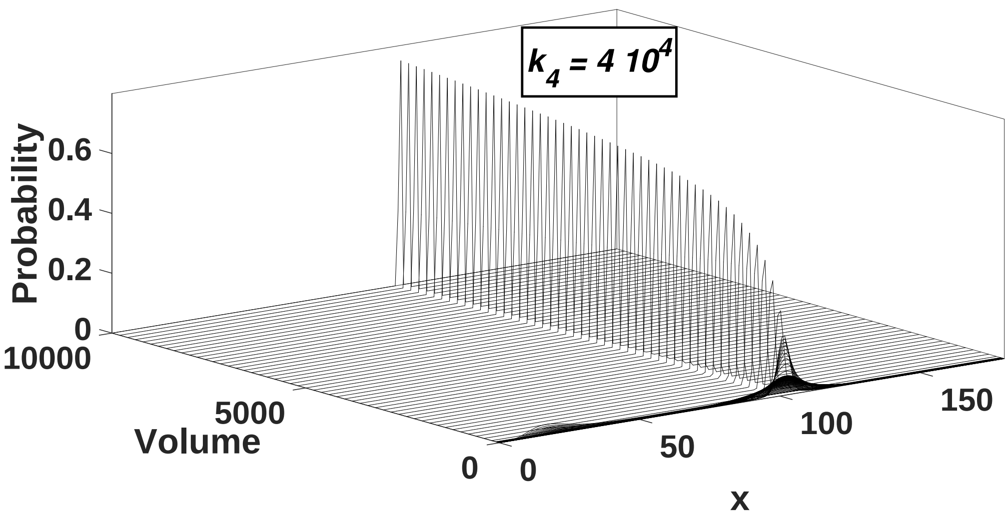
Figure 4.2: The probability distributions are shown for different volumes, for $k_4 = 3.5\ 10^4$ (figure 4.2c). The system starts with bimodal distributions at small volumes (figure 4.2a). Bimodality collapses as the system size increases. The stochastic model attains only one of the two deterministic solutions as the volumes increases (figure 4.2b). The solid vertical lines represent the stable deterministic solutions. Figure 4.2c has 220 plotted probability distributions computed with ZI closure scheme.



(a)



(b)



(c)

Figure 4.3: Similar results to figure 4.2, however here $k_4 = 4.0\ 10^4$. Bimodality collapses as the system size increases as well. However in this case, the stochastic model attains a different deterministic attractor as the volumes increases (figure 4.3b) compared to figure 4.2. The solid vertical lines represent the stable deterministic solutions and figure 4.3c has 220 plotted probability distributions computed with ZI closure scheme.

As the volume increases, the model's bimodality disappears, with only one peak remaining significant. The disappearance of bimodality occurs gradually with the system size, as one of the peaks becomes progressively smaller. Even though the less dominant peak might still exist in finite volumes, peak sizes fall below computer round-off errors. We speculate then that at the thermodynamic limit only one peak survives. Not unexpectedly, as the volume increases, probability distribution standard deviations decrease, resulting in a delta function centered at one of the ODE solutions (figures 4.2b & 4.3b). In other words for large system sizes, the stochastic model reaches only one solution whereas the ODE model can have three solutions, two stable and one unstable.

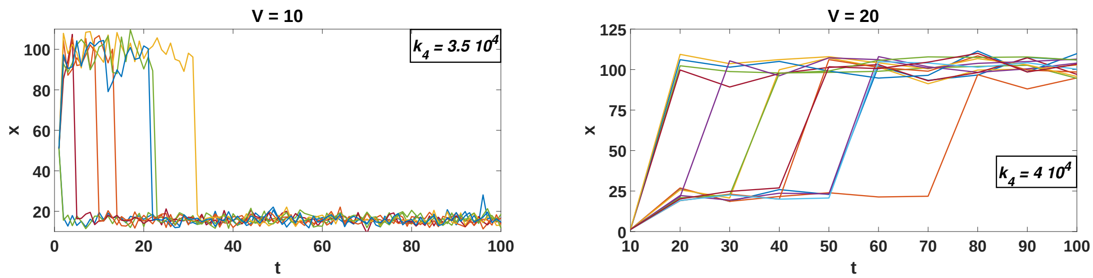
Because of system size models previously unattainable with stochastic simulations and because of the unexpected results, we have taken pains in ensuring the convergence and accuracy of the ZI closure scheme results presented herein. The reader is directed to the Appendix A.1 for more information about the accuracy of ZI closure results. There are three ways to increase confidence in ZI closure results:

First, in principle, for stochastic systems that satisfy the ergodic hypothesis, the CME solution does not depend on initial conditions. In practice, we have validated this hypothesis by determining CME solutions with the ZI closure scheme, numerically starting with various initial probability distribution solutions (prescribed in the initial values of the Lagrange multipliers in the numerical scheme).

Second, ZI closure solutions of the CME are verified by the analytical expressions available for the Schlögl model [8,68]. We note that even though the analytical solution is available, the numerical implementation becomes impractical for large volumes ($V > 40$ in this study) because of numerical range limitations in the calculation of the probability at $X = 0$ (Appendix A.1).

Finally, ZI closure scheme solutions is verified by comparison to the probability distribution obtained from stochastic simulation algorithm (SSA) simulations [9].

In light of the bimodality collapse results, it is interesting to conduct SSA simulations with trajectories that start at one of the ODE solutions. All trajectories moving to the attractor that is attained by ZI closure scheme, regardless of initial conditions. Surprisingly, system trajectories quickly transition to the stationary attractor, even if they are initiated at the second attractor (figure 4.4). This behavior is counterintuitive, especially in light of the deterministic model solutions, where the steady state is attained that is closest to



(a) (b)

Figure 4.4: Sample SSA trajectories for two different kinetic constants (figure 4.4a: $k_4 = 3.5\ 10^4$ and figure 4.4b: $k_4 = 4\ 10^4$). Even though trajectories start from a different attractor, they transition to the stochastic stationary one. As in figure 4.1, x is the concentration X/V. For visual clarity, the figure presents only a random sample of 500,000 simulated trajectories, all of which behave in a similar fashion.

the initial conditions of a simulation.

A quantity that has been used to explain the collapse of bimodality in the Schlögl probability distribution for mesoscopic systems is the passage or transition time. Mean (first) passage times can be calculated in a single-step, one-dimensional Markov chain with a bimodal distribution. They represent the time that is required for the system to move from one peak to the other and can be calculated from the following formulas [8]:

$$T_{ac} = \sum_{x=a}^{c-1} \frac{1}{W_+(x)P(x)} \sum_{y=0}^{x} P(y) \qquad T_{ca} = \sum_{x=a+1}^{c} \frac{1}{W_-(x)P(x)} \sum_{y=x}^{\infty} P(y), \tag{4.3}$$

where, T_{ac} is the mean passage time for moving from the left peak (a) to the right (c) and T_{ca} the time for moving from the right to the left, at steady state. $W_+(X) = \frac{k_2}{V}(X)(X - 1) + k_4 V$ and $W_-(X) = \frac{k_1}{V^2} X(X-1)(X-2) + k_3 X$ are the transition rates.

By calculating the mean passage time for different volumes (figure 4.5), it is observed that the time moving from the right peak to the left one over the time that it takes from left to right increases exponentially. The results support previous findings [57]. Even though a second peak might exist in higher volumes, it is significantly less important than the dominant one. Again, this implies that only one peak may be recovered at the thermodynamic limit.

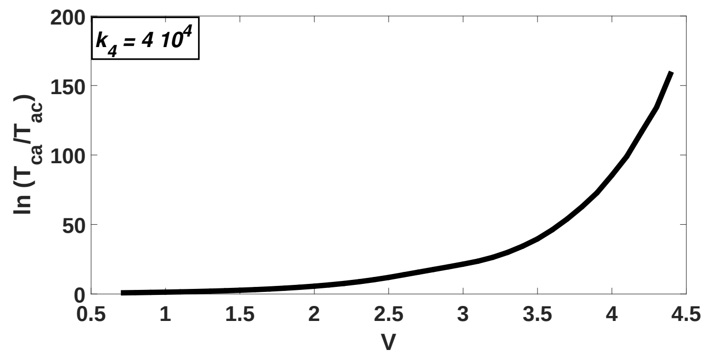


Figure 4.5: The figure shows the ratio of the time required to transition from the right peak to the left one over the time to move from the left to the right. In this figure, results for only one kinetic constant value ($k_4 = 4\ 10^4$) are presented.

4.2.2 Entropy in NESS

ZI closure postulates that a stochastic reaction system reaches the steady state probability distribution that maximizes the system's entropy, calculated as $H = -\sum_x P(x) \ln P(x)$, where the sum extends over all possible state space points x_i, assuming their number is finite. Studying how the maximum entropy of NESS (Non-Equilibrium Steady State) systems changes with the kinetic constants and the volume provides unique insights about the system. As mentioned in section 4.1, stationary probability distribution and NESS are two different words describing the same element, the stationary solution. In this section, the intention is to shift the attention to the behavior of the system as a whole at the steady state and not just describe the form of the solution. Thus, the term 'NESS' will be used to describe the steady state solution of the reaction network.

In figure 4.6, the maximum entropy of different computed NESS is plotted as a function of k_4 at volume $V = 1$. Each NESS represents the stationary probability distribution of a system. As such, a NESS can have only one value for system entropy. By the term 'maximum entropy of NESS', the author refers to the value of the system entropy at steady state and want to emphasize that this value is the maximum entropic value that the system can attain, for certain values of the kinetic constants.

At low kinetic constant values, the system is located at the unimodal region (figure 4.6A). At this region, the maximum entropy increases with increasing k_4. After a certain value of k_4, the slope of the graph changes and the system becomes bimodal (figure 4.6B). The increase of the entropy slope with k_4 is more pronounced in this bimodal regime. At a critical k_4 value, NESS maximum entropy reaches a maximum. At this critical value the peaks have the same probability (figure 4.6C). Then, the probability distribution transitions to a bimodal shape with a different attractor more profound (figure 4.6D). The entropy of the system begins to decrease. At a certain value of k_4, the probability distribution becomes again unimodal (figure 4.6E) resulting in yet another slope change.

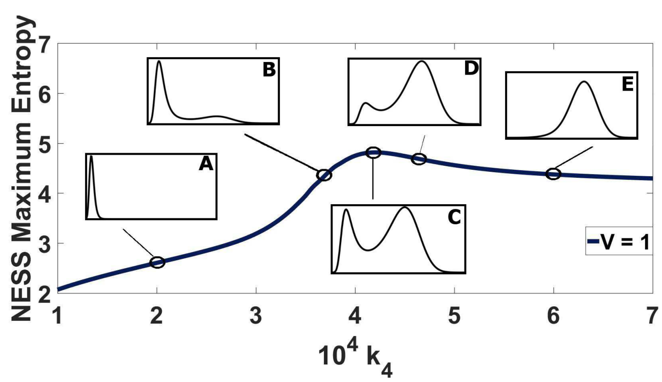


Figure 4.6: NESS entropy for varying kinetic constant (k_4) values at $V = 1$. Example stationary probability distributions are depicted, as insets, to indicate the five distinct trends of NESS entropy with k_4. The corresponding kinetic values are: $k_4 = 2\ 10^4$ (figure 4.6A), $k_4 = 3.7\ 10^4$ (figure 4.6B), $k_4 = 4.21\ 10^4$ (figure 4.6C), $k_4 = 4.5\ 10^4$ (figure 4.6D) and $k_4 = 6\ 10^4$ (figure 4.6E).

Changes of the maximum entropy of the NESS reflect changes of the shape of the distribution. As the kinetic constant increases the left peak becomes progressively smaller and the right one larger. Thus, the variance of the distribution changes as k_4 increases.

At the first unimodal region (figure 4.6A), the variance increases with k_4. As a result, the maximum entropy increases as well (figure 4.6). Then, the system becomes bimodal (figure 4.6B), the variance then has a different slope at the point that the system becomes bimodal. That is the point that maximum entropy of NESS also has a change in slope for the first time (figure 4.6). As the right peak becomes larger the variance and the maximum entropy become larger as well. At figure 4.6C, the two peaks are expressed equally and the maximum entropy reaches a maximum.

Similar behavior is observed as we start from high k_4 values (figure 4.6E), right side from point C of figure 4.6. The system is unimodal and as k_4 is decreasing we transition to unimodality (figure 4.6D). Thus, the variance and maximum entropy increase with a decrease in k_4. At the point that the system becomes bimodal for the first time the slope changes.

Figure 4.7, present the same graph as figure 4.6 for four different system sizes. In fact, hundreds of NESS probability distributions are computed, varying volume by 50, and the corresponding entropies plotted in this graph. As the volume increases the region between the two slope changes becomes narrower. This region also contains the maximum point of the graph. As discussed earlier, the aforementioned region is the bimodal region of the NESS. Thus, as the volume increases the bimodal region shrinks. For larger volumes, there is a wider range of kinetic constant values that had a unimodal stationary probability distribution.

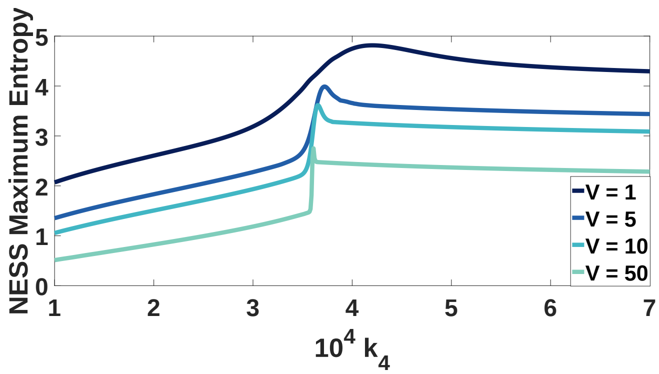


Figure 4.7: The maximum entropy of different NESS probability distributions plotted against the kinetic constant k_4, for four different volumes V. Each volume line includes more than 600 points, with each one of them corresponding to a different NESS. With so many points there is the appearance of a continuous line for each V.

In other words, there is a range of k_4 values for which the probability maintains its bimodal character. This range narrows for higher volumes. For example, at system size $V = 10$, this region is a much smaller fraction of the k_4 range compared to $V = 1$. The entropy changes become more pronounced and sharper for the larger size. This leads us to speculate that at the thermodynamic limit, the bimodal region will shrink that much that will contain only a single critical k_4 value. At this point, the probability distribution exhibits two equally weighted distinct peaks whose integral sums up to 1, with a total entropy of $-\ln(0.5)$.

Since only one critical k_4 value gives a bimodal distribution at the thermodynamic limit, every other k_4 value will give a unimodal distribution. For smaller k_4 values than the critical, the left attractor will be expressed and, for larger values than the critical, the right attractor will be expressed. Thus, the probability distribution would transition from a single delta function to two equal-sized delta-like functions positioned at the two stable ODE solutions at the critical k_4, and then back to a different single delta function. The fact that the distribution at thermodynamic limit will be delta function was discussed at section 4.2.1 based on the evidence of figures 4.2 & 4.3. We can only speculate this thermodynamic limit behavior, because ZI closure scheme is a numerical method. In other words, at the critical kinetic constant values the system would move through a phase transition. To the author's knowledge, this is the first time such results are presented.

4.2.3 Stochastic NESS Jacobian Eigenvalues of Mesoscopic Systems

We have discussed how the eigenvalues of a stochastic system can be calculated based on the Jacobian around steady state [40]. Chapter 5 presents in mathematical detail these calculations. The purpose of this section is to explore the dependence of eigenvalues on the size of the system.

When the system is bimodal, absolute eigenvalues are smaller, which implies slower system dynamics. This makes sense: for unimodal systems there is the one timescale of the system exploring its single attractor. On the other hand, for bimodal systems, the system dynamics exhibits an additional timescale for switching between two attractors. In figure 4.8, the dominant eigenvalue of the steady state Jacobian is presented as a function of the volume for three values of k_4. For all three values of k_4, the CME admits a bimodal probability distribution at small volumes. All eigenvalues are negative indicating stable NESS.

As the volume increases, the eigenvalues of systems with $k_4 = 3.5\,10^4$ and $k_4 = 4\,10^4$ decrease. This indicates a faster dynamical response to small perturbations away from the NESS. Equivalently, and in accord with the fluctuation-dissipation theorem, these eigenvalues mean that fluctuation correlations of the first probability distribution moment decay faster for larger volumes.

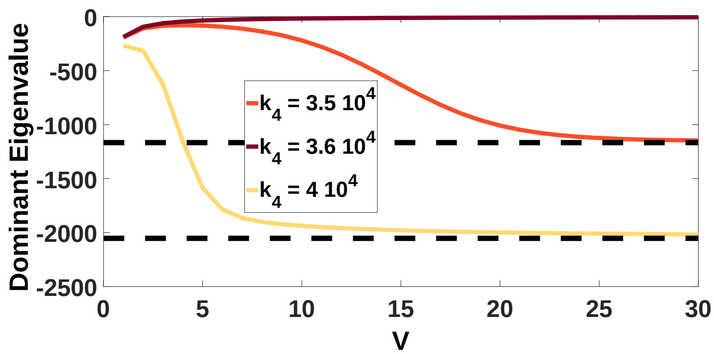


Figure 4.8: The dominant eigenvalue (corresponding to the first probability distribution moment) of the stochastic system with respect to volume V. Three different values of k_4 are presented. The black dashed lines represent the deterministic eigenvalue of the corresponding attractor. Similar to figure 4.7 each solid line corresponds to more than 700 NESS systems. For visualizing purposes, all the points were connected to a continuous line.

The dominant eigenvalues of the stochastic steady state Jacobian asymptotically tend to the eigenvalue of one of the two stable ODE solutions. This means that bimodality collapses and the stochastic system solution resembles one of the ODE stable solutions, in accord with figures 4.2 & 4.3 results.

The behavior is different for $k_4 = 3.6\ 10^4$. This value corresponds to the region of the critical k_4 value, as discussed in figure 4.7. The eigenvalue increases with increasing system size and asymptotically tends to zero. At this value of k_4, the probability distribution is bimodal with two sharp peaks even at very large volumes. The transition times between peaks diverge to infinity, and fluctuation correlations cease to decay in finite time scales.

4.3 Discussion

The differences between deterministic and stochastic models for chemical reaction networks have long been the subject of numerous investigations [55–57]. In this chapter, results are presented that highlight such differences, indicating the collapse of bimodality at the mesoscopic area of the Schlögl model. Whereas the ODE model gives two stable solutions, the stochastic only gives one. The size of the system is an important parameter. For small systems, the stochastic model produces a bimodal stationary distribution with peaks near the ODE solutions. As the system size increases, the stochastic model gradually prefers one of the two distribution peaks. This behavior is observed with ZI closure scheme solutions. The question then may arise: what drives a mesoscopic system to only one of the two attractors? Especially, when the initial condition is the other attractor and both of them are stable solutions of the ODE model.

Numerous explanations have been provided for the discrepancies, including ones based on the mean transition times between peak, and most notably the entropy production rate [66, 68]. In fact, numerous efforts have attempted, and ultimately failed to establish whether a criterion for non-equilibrium steady states (NESS) can exist based on the rate of entropy production [44, 68–71].

In this study, the stationary probability distributions are reached when the entropy of the system is maximum. We stress that the ZI closure scheme postulates a maximum entropy for every NESS and then computes an accurate probability distribution. In other words, the accurate solution of the CME obtained by the ZI closure scheme is the probability distribution with maximum entropy, for any given set of parameter values. We wonder whether this may be a general criterion for establishing NESS of chemical reacting systems. This is also a finding that warrants further study.

Note that the total entropy of the system and the two reservoirs increases with time. Given a non-zero difference between the chemical potentials of the two reservoirs, there is mass transfer from the high to the low chemical potential reservoir. This is overall an irreversible process with positive entropy production. Yet, within the system, a stationary probability distribution is attained, which implies that the entropy of the system contained within volume V is constant. Additionally, for open, isothermal, isochoric systems, the relevant equilibrium thermodynamic property is the Helmholtz free energy. Herein the defined Schlögl model is void of enthalpic constraints. Thus a minimum Helmholtz free energy is equivalent to a maximum entropy.

4.4 Summary

In this chapter solutions of the CME for Schlögl model systems with microscopic, mesoscopic and macroscopic sizes are presented. The results cover a wide variety of system sizes, previously unattainable with traditional methods (e.g. SSA, analytical solution). This is possible now with the employment of ZI closure scheme. As reported previously, for mesoscopic systems, the exhibited bimodality collapses and only one of the attractors becomes dominant with probability 1.

It is revealed that the range of kinetic constant values where bimodality survives in larger systems diminishes in size. Presumably, there is a single value of k_4 where bimodality survives at the thermodynamic limit. This is only speculative because, although the ZI closure scheme affords the investigation of system sizes that remained inaccessible in previous studies, it is also still limited to finite volumes.

It is also found that the dominant eigenvalue of the stochastic system asymptotically reaches the eigenvalue of the corresponding deterministic solution as the volume increases. This implies that at the macroscopic limit the NESS probability distribution not only replicates one of the deterministic solutions but also fully regains the dynamic behavior of it as well. Thus at the thermodynamic limit, the bimodality collapses and the remaining unimodal distribution fully represents the behavior of one of the deterministic solutions.

The ZI closure scheme postulates the maximum entropy stationary probability distribution in order to numerically close the probability distribution moment equations.

Chapter 5

A Linearization Method for Probability Moment Equations *

5.1 Introduction

Recently stochastic reaction models have been the focus of numerous studies because they describe important aspects of biological processes more accurately than deterministic reaction kinetic models [7, 36, 47]. In biomolecular systems, the number of molecules of reacting species is often very small. Stochasticity then impacts the observed cellular phenotypes, requiring modeling formalisms that consider the inherent variability in biological systems [5].

The prevalent techniques to model stochastic reaction networks have largely relied on Gillespie's Stochastic Simulation Algorithms (SSA) [8, 9, 12, 31, 72, 73]. The SSA is an established modeling formalism that accurately captures the discrete and probabilistic nature of biomolecular interactions. However, the SSA and its numerous variants suffer from significant computational inefficiencies, principally because each reaction event is

*Reprinted (adapted) with permission from Michail Vlysidis and Yiannis N. Kaznessis, *Computers & Chemical Engineering*, **112**:1-5, 2018 [40].

explicitly accounted for in the model. The computational challenge becomes insurmountable when there are system reactions with multiple, disparate time scales. The SSA will then spend inordinate amounts of computer time, simulating fast reaction events, in order to properly sample slower reaction events [72].

Alternative methods for modeling stochastic reaction networks that rely on the moments of the probability distribution of these networks have been developed [1, 35, 39]. Assuming the entropy of the network is maximum, we developed a closure scheme for probability moment equations (ZI closure scheme). This scheme can be employed to quickly and accurately compute the steady state of non-linear reaction networks.

It is essential for understanding the behavior of stochastic networks to be able to capture the time dynamic evolution of them. Unfortunately, current numerical methods, including ZI closure scheme, can succeed on accurately capturing the stationary behaviors but fail on calculating time transient behaviors. Thus, there is a need for a new approach that can predict the whole dynamics of stochastic networks.

Linearization techniques have been developed for over six decades to solve non-linear initial value problems in chemical kinetics [74–76]. Such techniques have also been applied to a plethora of other fields besides reaction kinetics [77–81]. Linearization of non-linear reaction rate equations requires knowledge of their steady-state solution. Using this information and an initial condition, linearized equations can be solved to compute time-dependent behaviors [82, 83].

Based on this moment closure steady state approach, herein a linearization method for probability moment equations to calculate the time-transient behavior of stochastic reaction networks close to steady state is proposed. This new method can be used to model reaction networks away from the thermodynamic limit, including biomolecular and metabolic reaction networks in biological systems.

In the following section, the theoretical underpinnings of the linearization method are presented.

5.2 Theory

The background theory has been discussed in detail in Chapters 1 & 2. A complete model for describing the time evolution of the probability distribution of a system with chemical reactions is the Chemical Master Equation (CME) [5, 8]. For all its elegance, the CME is not often used in modeling reaction systems because it is not tractable for any but the simplest of reacting systems [35].

5.2.1 Moment Equations

An alternative method for computing probability distributions relies on calculating probability moments. There exist different types of moments, such as central, factorial, or polynomial ones, all equally valid for describing the probability distribution [7, 22, 25]. The derivations in this chapter apply to any kind of probability moments.

For a chemical reaction network with N reactants and products, the state of the network is described by the N-dimensional vector $\boldsymbol{X} = (X_1 \ldots X_N)$ that contains the number of molecules of each component (all reactants and products). If the functional form of the moment μ_i is denoted as f_{μ_i}, then the moment itself is defined as follows:

$$\mu_i = \sum_{\Omega} f_{\mu_i}(\boldsymbol{X}) \cdot P(\boldsymbol{X}; t) \tag{5.1}$$

where, $P(\boldsymbol{X}; t)$ is the probability for the system to be at state $\boldsymbol{X}$ at time t and Ω correspond to the N-dimensional state space.

With no loss of information, CMEs can be transformed into moment equations through a Z-transform [22]. The resulting moment equations are a set of coupled differential equations. The most general form for a system with elementary reaction rates is:

$$\frac{\partial \boldsymbol{\mu}}{\partial t} = A \cdot \boldsymbol{\mu} + A' \cdot \boldsymbol{\mu'} + \boldsymbol{\mu_c} \tag{5.2}$$

where $\boldsymbol{\mu}$ is the lower-order moment vector (not including the zero-order moment, which is always equal to 1), $\boldsymbol{\mu'}$ is the higher-order moment vector, $\boldsymbol{\mu_c}$ is a vector of constants, and t represents time.

In most cases, vectors $\boldsymbol{\mu}$ and $\boldsymbol{\mu'}$ have different dimensions (Chapter 1). For non-linear reaction networks, the $\boldsymbol{\mu'}$ vector is nonempty. It is then apparent that the lower-order moments depend on the higher-order ones, and the set of equations cannot be closed.

5.3 Calculation/Linearization

Moment equations can be linearized around the stationary probability distribution by computing the steady state Jacobian (J_{ss}). By carrying out a Taylor expansion of the right side of equation (5.2) around the steady state, we can use the following approximation:

$$\frac{\partial \boldsymbol{\mu}}{\partial t} \approx \left.\frac{\partial \boldsymbol{\mu}}{\partial t}\right|_{ss} + J_{ss} \left(\boldsymbol{\mu} - \boldsymbol{\mu_{ss}}\right) \tag{5.3}$$

where the subscript ss stands for steady state, i.e. $\boldsymbol{\mu_{ss}}$ denotes the stationary probability moments. When evaluated at steady state the time derivative is zero ($\left.\frac{\partial \boldsymbol{\mu}}{\partial t}\right|_{ss} = \mathbf{0}$). As a result equation (5.3) is simplified to

$$\frac{\partial \overline{\boldsymbol{\mu}}}{\partial t} = J_{ss} \overline{\boldsymbol{\mu}} \tag{5.4}$$

with $\overline{\mu} = (\boldsymbol{\mu} - \boldsymbol{\mu_{ss}})$.

Section 5.3.1 provides details on the calculation of steady state Jacobian, based on the knowledge of stationary probability moments.

Equation (5.4) has a known analytical solution [47], based on the eigenvalues ($\boldsymbol{\lambda_e}$) and eigenvectors ($\boldsymbol{v}$) of the Jacobian matrix J_{ss}. Through an eigenvalue analysis [30,47], equation (5.4) can be transformed into:

$$\boldsymbol{\mu}(\boldsymbol{t}) = \boldsymbol{v}\exp(\boldsymbol{\lambda_e} t)\boldsymbol{v^{-1}}\left[\boldsymbol{\mu}(\boldsymbol{0}) - \boldsymbol{\mu_{ss}}\right] + \boldsymbol{\mu_{ss}} \tag{5.5}$$

where t is the time that the moment vector $\boldsymbol{\mu}$ is evaluated and $\boldsymbol{v^{-1}}$ is the inverse of the eigenvector matrix $\boldsymbol{v}$.

With this method, one can calculate probability moments at any time, with only the knowledge of the stationary solution (i.e. the Jacobian matrix around the steady state) and an initial condition ($\boldsymbol{\mu}(\boldsymbol{0})$) of the moments.

5.3.1 Steady State Jacobian Matrix

By definition the Jacobian matrix is given by:

$$J_{ss} = \left.\frac{\partial\left[\frac{\partial\boldsymbol{\mu}}{\partial t}\right]}{\partial\boldsymbol{\mu}}\right|_{ss} = A + A'\left.\frac{\partial\boldsymbol{\mu'}}{\partial\boldsymbol{\mu}}\right|_{ss} \tag{5.6}$$

Smadbeck & Kaznessis [28] developed ZI closure scheme method to connect lower and higher-order moments with the Lagrange multipliers ($\boldsymbol{\lambda}$) of the maximum entropy distribution. Thus, with the use of the chain rule for partial derivatives the following stands:

$$\left.\frac{\partial\boldsymbol{\mu'}}{\partial\boldsymbol{\mu}}\right|_{ss} = \left.\frac{\partial\boldsymbol{\mu'}}{\partial\boldsymbol{\lambda}}\right|_{ss}\left.\frac{\partial\boldsymbol{\lambda}}{\partial\boldsymbol{\mu}}\right|_{ss} = \left.\frac{\partial\boldsymbol{\mu'}}{\partial\boldsymbol{\lambda}}\right|_{ss}\left(\left.\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\lambda}}\right|_{ss}\right)^{-1} \tag{5.7}$$

The Jacobian matrix can be then expressed [48] as:

$$J_{ss} = A + A'\left.\frac{\partial\boldsymbol{\mu'}}{\partial\boldsymbol{\lambda}}\right|_{ss}\left(\left.\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\lambda}}\right|_{ss}\right)^{-1} \tag{5.8}$$

In order to calculate the steady state Jacobian one needs to calculate matrices $\left.\frac{\partial\boldsymbol{\mu'}}{\partial\boldsymbol{\lambda}}\right|_{ss}$ and $\left.\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\lambda}}\right|_{ss}$.

Assuming the entropy of a reacting system is maximum, the probability distribution is calculated using constrained optimization, where Lagrange multipliers are related to the probability distribution with the following equation:

$$P(\boldsymbol{X}) = \exp\left(-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right) \tag{5.9}$$

where M is the size of matrix $\boldsymbol{\mu}$ and λ_j the j^{th} Lagrange multiplier. Since $\sum_\Omega P(\boldsymbol{X}) = 1$ for any probability distribution, the zero$^{\text{th}}$ Lagrange multiplier, λ_0, depends on the rest of Lagrange multipliers through the expression:

$$\begin{aligned} \sum_\Omega P(\boldsymbol{X}) &= 1 \\ \sum_\Omega \left[\exp\left(-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right] &= 1 \\ \exp(-\lambda_0) \sum_\Omega \left[\exp\left(-\sum_{j=1}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right] &= 1 \\ \lambda_0 = \log\left\{\sum_\Omega \left[\exp\left(-\sum_{j=1}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]\right\} & \end{aligned} \tag{5.10}$$

Based on equations (5.1) and (5.9), the probability moments are related with the Lagrange multipliers:

$$\begin{aligned} \mu_i &= \exp(-\lambda_0) \sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right) \\ &= \frac{\sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right)}{\sum_\Omega \left[\exp\left(-\sum_{j=1}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]} \end{aligned} \tag{5.11}$$

The above equation can be used to calculate matrices $\left.\frac{\partial \boldsymbol{\mu}'}{\partial \boldsymbol{\lambda}}\right|_{ss}$, $\left.\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\lambda}}\right|_{ss}$ and thus the steady state Jacobian:

$$\frac{\partial \mu_i}{\partial \lambda_{m_j}} = \frac{\frac{\partial}{\partial \lambda_{m_j}} \sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)}{\sum_\Omega \left[\exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]}$$
$$-\frac{\sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)}{\sum_\Omega \left[\exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]} \frac{\frac{\partial}{\partial \lambda_{m_j}} \sum_\Omega \left[\exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]}{\sum_\Omega \left[\exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]} \quad (5.12)$$

It is true that:

$$\frac{\partial}{\partial \lambda_{m_j}} \sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)$$
$$= -\sum_\Omega f_{\mu_i}(\boldsymbol{X}) \cdot f_{\mu_j}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right) \quad (5.13)$$

and:

$$\frac{\partial}{\partial \lambda_{m_j}} \sum_\Omega \left[\exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]$$
$$= -\sum_\Omega f_{\mu_j}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right) \quad (5.14)$$

By combining equations (5.12), (5.13), (5.14) and the definition of moments (equation (5.11)), the following holds:

$$\frac{\partial \mu_i}{\partial \lambda_{m_j}} = -\mu_{i,j} + \mu_i \cdot \mu_j \quad (5.15)$$

where $\mu_{i,j}$ is the combined moment of moments f_{μ_i} and f_{μ_j}, given by:

$$\mu_{i,j} = \frac{\sum_\Omega f_{\mu_i}(\boldsymbol{X}) f_{\mu_j}(\boldsymbol{X}) \exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)}{\sum_\Omega \left[\exp\left(-\sum_{j=1}^M \lambda_j f_{\mu_j}(\boldsymbol{X})\right)\right]} \quad (5.16)$$

In summary, the Jacobian matrix is given by:

$$J_{ss} = A + A' \left.\frac{\partial \boldsymbol{\mu'}}{\partial \boldsymbol{\lambda}}\right|_{ss} \left(\left.\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\lambda}}\right|_{ss}\right)^{-1} \quad (5.17)$$

where $\frac{\partial \mu_i}{\partial \lambda_{m_j}} = -\mu_{i,j} + \mu_i \cdot \mu_j$ and $\mu_{i,j} = \sum_\Omega f_{\mu_i}(\boldsymbol{X}) \cdot f_{\mu_j}(\boldsymbol{X}) \cdot P(\boldsymbol{X};t)$.

5.4 Results

Three different reaction networks are employed in order to demonstrate the proposed linearization method. The first network is the Schlögl model [46], the second one is the Michaelis-Menten model [1, 22] and the third one is Wilhelm's bistable system [49] (table 5.1). In order to assess the accuracy and computational efficiency of this method, the results are compared with SSA results using the same initial condition and kinetic constants.

Table 5.1: The table shows the reaction networks for the Schlögl, Michaelis-Menten and Wilhelm's bistable models. The kinetic constants values and initial condition for each system are also presented. For the third order reactions the kinetic constants are in $(molecules^2 \cdot s)^{-1}$ units, for the second order in $(molecules \cdot s)^{-1}$, for the first order in s^{-1} and for the zeroth order in $molecules \cdot s^{-1}$. Initial conditions are in $molecule$ units.

Models	*Reactions*	*Kinetic constants*	*Initial Moments*
Schlögl	$3X \xrightarrow{k_1} 2X$	$k_1 = 7.5\ 10^{-5}$	$X_0 = 25$ (unimodal) $X_0 = 50$ (bimodal)
	$2X \xrightarrow{k_2} 3X$	$k_2 = 0.0075$	
	$X \xrightarrow{k_3} \emptyset$	$k_3 = 0.175$	
	$\emptyset \xrightarrow{k_4} X$	$k_4 = 2$ (unimodal) $k_4 = 1.1$ (bimodal)	
Michaelis-Menten	$S + E \xrightarrow{k_1} S:E$	$k_1 = 0.1$	$S_0 = 100$ $E_0 = 50$ $S:E_0 = 0$ $P_0 = 0$
	$S:E \xrightarrow{k_2} S + E$	$k_2 = 0.1$	
	$S:E \xrightarrow{k_3} P + E$	$k_3 = 1$	
	$P \xrightarrow{k_4} S$	$k_4 = 1$	
Wilhelm	$Y \xrightarrow{k_1} 2X$	$k_1 = 35$	$X_0 = 10$ $Y_0 = 10$
	$2X \xrightarrow{k_2} X + Y$	$k_2 = 1$	
	$X + Y \xrightarrow{k_3} Y$	$k_3 = 1$	
	$X \xrightarrow{k_4} \emptyset$	$k_4 = 9.74$	
	$\emptyset \xrightarrow{k_5} X$	$k_5 = 30$	

For each system, the stationary probability distribution was calculated with ZI closure scheme [28]. This was then used for the calculation of the steady state Jacobian (J_{ss}) as described in section 5.3.1. From the Jacobian matrix, eigenvalues ($\boldsymbol{\lambda_e}$) and eigenvectors ($\boldsymbol{v}$) were determined with Matlab 2016b. Finally, equation (5.5) is employed with a set of initial conditions, in order to calculate the transient behavior of the probability distribution moments.

The Schlögl model is one of the simplest theoretical single-component systems that can exhibit bistability [49]. Depending on the kinetic constants, the Schlögl model either has a unimodal or bimodal stationary probability distribution. Results for both unimodal and bimodal solutions are presented. In both cases, the method accurately calculates the time evolution of the lower-order moments of the probability distribution (figures 5.1 & 5.2). Based on the results of chapter 4, we have found that up to ten moments are necessary to accurately capture the shape of bimodal probability distributions. There is an increasing interest in both synthetic and natural biological systems exhibiting bistability [61–63], and the Schlögl model is often used as a simple prototypical model.

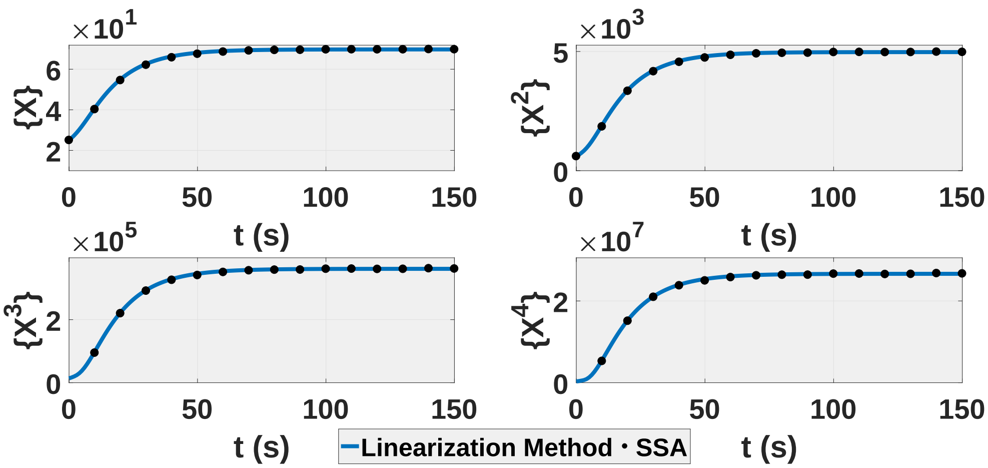


Figure 5.1: Time evolution of the probability moments for the unimodal region of the Schlögl model. Solid line represents results calculated based on equation 5.4 and dots represents results calculated with SSA. Results are presented for the first four moments of the model.

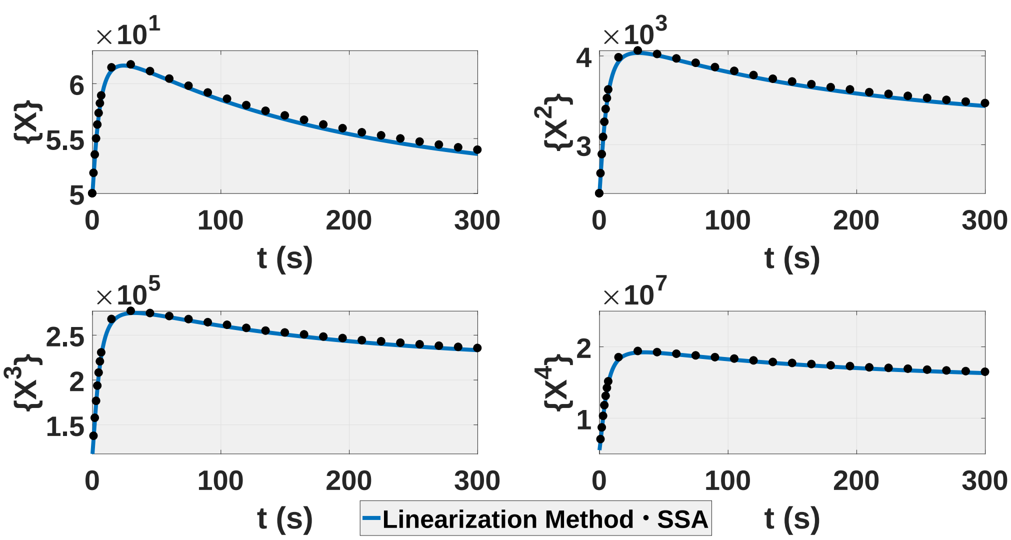


Figure 5.2: The figures show the time evolution of the first four moments of the Schlögl model for the bimodal region, similar to figure 5.1. Solid lines are calculated with equation 5.4 and dots with the SSA.

Wilhelm's network is another network with bistable behavior. It was first introduced as one of the smallest known bistable chemical reaction systems and is a two-component non-linear reaction network [49]. The moments of this network have a complex behavior with both minima and maxima appearing for the same moment (figure 5.3).

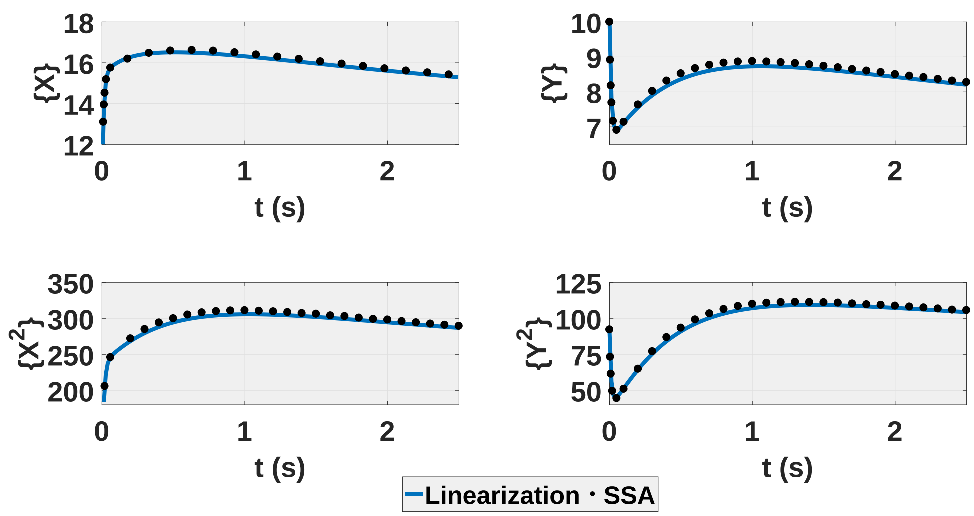


Figure 5.3: Time evolution of the first five moments of the Wilhelm's bistable model, similar to figures 5.1, 5.2 & 5.4. Solid lines are calculated with equation 5.4 and dots with the SSA.

Results for a Michaelis-Menten system are presented in figure 5.4. Michaelis-Menten is also a two-component non-linear reaction network. The network is a closed system with a unimodal stationary probability distribution. Michaelis-Menten is a well studied network and one of the most popular enzyme kinetics models [50, 51]. It is commonly used to describe enzymatic catalysis [84].

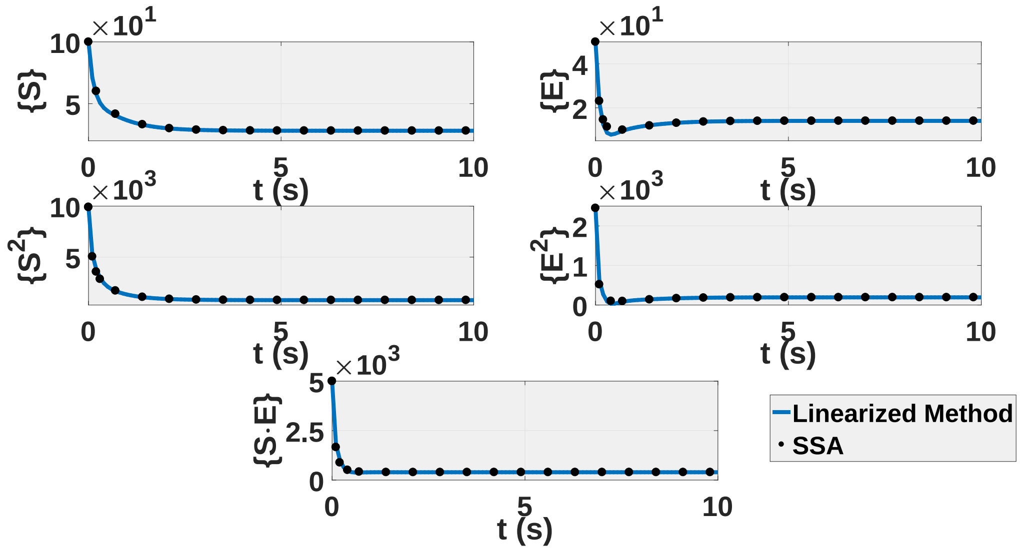


Figure 5.4: Time evolution of the first five moments of the Michaelis-Menten model, similar to figures 5.1 & 5.2. Solid lines are calculated with equation 5.4 and dots with the SSA.

The time required to reach steady state is similar for all the moment orders of each example. The dynamics of the stochastic systems are then equally influenced by all the probability moments. The error between the calculated moment values and SSA moments is listed in table 5.2. For each case, 500,000 SSA trajectories were simulated, and the percent difference between the two methods was calculated for each time point until steady state was reached. Table 5.2 presents the average of the error over all the time points. In each case, the average error is relatively small (below 3 %).

The computational cost of the linearization approach is insignificant, since it produces an analytical solution for the moments (equation (5.5)). Assuming the stationary moments are known, the method requires the construction of the stationary Jacobian and the calculation of its eigenvalues and eigenvectors, all of which are tasks with minimal computational cost. In comparison, the computational cost of 500,000 SSA trajectories varied between 30 and 50 CPU hours.

For all the presented examples, the initial condition was a delta function; the initial condition values are listed in table 5.1. The shape of the initial distribution is different than the stationary distribution of all systems. As a result, the initial conditions are significantly far from the stationary moments, with the first moment being the only exception. The proposed method is still able to accurately predict the time evolution of the moments, as long as the initial condition of the first moment is relatively close to the stationary value.

Table 5.2: The table presents results for the Schlögl (both unimodal and bimodal), Michaelis-Menten and Wilhelm's bistable models. The average error between the moments calculated with the proposed method against SSA. For each time point the error was calculated as the normalized percentage difference between the two methods. The table presents the average error among all the time points of a moment until steady state is reached. For the Schlögl model the average error is presented for the first four moments. For the other two networks the error of the first five moments is presented.

Unimodal ***Schlögl***		*Bimodal* ***Schlögl***	
	Average error (%)		*Average error (%)*
$\{X\}$	0.07	$\{X\}$	0.17
$\{X^2\}$	0.12	$\{X^2\}$	0.26
$\{X^3\}$	0.17	$\{X^3\}$	0.31
$\{X^4\}$	0.23	$\{X^4\}$	0.35
Michaelis-Menten		***Wilhelm***	
	Average error (%)		*Average error (%)*
$\{S\}$	0.44	$\{X\}$	0.20
$\{E\}$	1.15	$\{Y\}$	0.34
$\{S^2\}$	0.42	$\{X^2\}$	0.36
$\{S \cdot E\}$	0.87	$\{X \cdot Y\}$	0.51
$\{E^2\}$	3.01	$\{Y^2\}$	0.49

5.5 Discussion

Linearization methods are numerical approximations, useful for calculating time transient behaviors based on the knowledge of a system's steady state. Herein, a linearization technique for solving the probability moment equation of stochastic reaction networks was proposed.

This linearizes moment equations through a two-term Taylor expansion. The time behavior of probability moments is then calculated through the known analytical solution of the linearized moment equations. The solution requires the eigenvalues and eigenvectors of the Jacobian matrix around the steady state solution. Thus, the calculation of the stationary Jacobian for the stationary probability distribution is a necessary intermediate step, for which we implement with the ZI closure scheme.

Three different stochastic reaction networks are studied to demonstrate the method: a network with two components and a non-linear behavior (Michaelis-Menten), a network with a single component and bistable behavior (the Schlögl model), and a two-component network with a more complicated bistable behavior (the Wilhelm model). In all three cases, the calculated moment behavior matches those calculated with SSA.

Equation (5.4) is a linear approximation around the steady state of equation (5.2). As such, the accuracy of the method depends on how close the initial conditions are from the steady state solution. The closer the initial guess is to the steady state solution, the more accurately the method will predict the time evolution of the moments.

Herein, only the first two terms of the Taylor expansion are employed for the derivation of equation (5.5). As the higher-order terms of the Taylor expansion become significant, the accuracy of the method may decrease. Systems with limit cycles are a prime example where higher-order Taylor terms are important. In such cases, more terms of the Taylor expansion may be necessary. A new study of the effect of including more terms on the computational cost and accuracy of the method may be warranted.

Chapter 6

Solving Stochastic Reaction Networks with Maximum Entropy Lagrange Multipliers

The linearizion technique discussed in Chapter 5 is a numerical approximation of the moment equations. The accuracy of the method depends on the initial conditions and the method is unable to provide results for oscillatory systems. In this chapter, we discuss a different approach on solving stochastic reacting systems.

6.1 Introduction

When a reaction network evolves stochastically, the Chemical Master Equation (CME) can model the probability distribution of the system being at a particular state at time t [5, 6]. The historical difficulties in solving the CME are well documented in the literature [7, 33, 47, 85]. An established way to solve CME is to use Gillespie's Stochastic Simulation Algorithm (SSA), a kinetic Monte Carlo method [9]. Despite the accuracy of SSA, the method is a kinetic Monte Carlo algorithm and hence computationally expensive [31].

A proposed alternative to solve CME is by calculating the probability moments [5]. This approach is based on the concept that any probability distribution can be completely described by its moments [36]. Moments are expected values with the most well-known ones being the mean and the variance [7,22,25]. The starting point of moment equations is to rewrite the CME in terms of moments describing the master probability distribution [1, 25]. Instead of a master equation that governs the probability distribution in time, one can then write a set of differential equations evolving the expected values of the distribution [31, 32]. Yet, because the dynamics of lower-order moments depend on the higher ones for non-linear reaction networks, the system of ODEs needs to be closed or somehow truncated in order to be solved [25, 28].

To tackle the issue with the numerical difficulties of moment equations and calculate the time evolution of stochastic networks, we hereby propose for the first time the use of Lagrange multipliers [28, 39, 42]. The connection between moments and Lagrange multipliers relies on the maximum entropy principle, which states that if some features of a probability distribution are known *a priori*, then in order to estimate the full distribution we can look for the one with maximum uncertainty that also satisfies the given knowledge (unbiased estimate) [1,42, 86].

Therefore, using the definition of information entropy by Shannon [41], the system of moment equations can be transformed into a system of equations that depend only on the Lagrange multipliers. We refer to the new system as ”Lagrange multiplier equations (LMEs)”. The LMEs set, unlike moment equations, is a closed system, i.e. it has the same number of unknowns with equations. Thus, an intial value problem numerical technique, like Runge-Kutta [87], can be used to solve the system for multiple times. Most importantly, the only assumption used is that the entropy of the system needs to be maximum at every time.

Herein, we unpack in full mathematical detail the generation of a general system of LMEs. We apply it to multiple non-linear reaction networks, including the bistable Schlögl model [46], the multistable Wilhelm model [49], the oscillatory Brusselator system [88,89] and the viral infection model [18,90]. After presenting a numerical procedure for the implementation of the method, we show results comparing the accuracy of LMEs solutions to SSA simulations. LMEs are able to solve reaction networks as accurate as SSA with significantly less computational time needed.

6.2 Theory

6.2.1 Lagrange Multiplier Equations

Connect Moments to Lagrange Multipliers

The general form of the moment equations is:

$$\frac{\partial \boldsymbol{\mu}}{\partial t} = A\boldsymbol{\mu} + A'\boldsymbol{\mu'} + \boldsymbol{\mu_c} \tag{6.1}$$

where $\boldsymbol{\mu}$ is the lower-order moment vector (not including the zero-order moment, which is always equal to 1), $\boldsymbol{\mu'}$ is the higher-order moment vector, $\boldsymbol{\mu_c}$ is a vector of constants and t represents time. A and A' are constant matrices that represent the linear and non-linear components of the network, respectively. In most cases, the vectors $\boldsymbol{\mu}$ and $\boldsymbol{\mu'}$ have different dimensions (Chapter 1). For non-linear reaction networks, the $\boldsymbol{\mu'}$ vector is nonempty.

For a chemical reaction network with N reactants and products, the moment μ_i is connected to the probability distribution P through the following expression:

$$\mu_i = \sum_{\Omega} f_{\mu_i}(\boldsymbol{X}) P(\boldsymbol{X}) \tag{6.2}$$

where f_{μ_i} is the functional form of the i^{th} moment μ_i. The number of molecules of each component is contained in the matrix $\boldsymbol{X} = (X_1 \ldots X_N)$ and Ω corresponds to the N-dimensional state space for all the possible values of $(X_1 \ldots X_N)$.

One can connect the probability distribution to the Lagrange multipliers $\boldsymbol{\lambda}$ by the maximizing information entropy [28]:

$$P(\boldsymbol{X}) = \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right] \tag{6.3}$$

where λ_j is the j^{th} Lagrange multiplier and M is the number of lower-order moments and also the size of vector $\boldsymbol{\mu}$. Thus, moments can be related to Lagrange multipliers, by combing equations (6.2) and (6.3):

$$\mu_i = \sum_{\Omega} \left\{ f_{\mu_i}(\boldsymbol{X}) \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right]\right\} \tag{6.4}$$

Time Derivative

In equation (6.3), only the Lagrange multipliers depend on time; the state space and the functional form of the moments are time independent. Hence, the time derivative of the probability distribution is:

$$\begin{aligned}
\frac{\partial P(\boldsymbol{X})}{\partial t} &= \frac{\partial \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right]}{\partial t} \\
&= \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right] \frac{\partial \left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right]}{\partial t} \\
&= -\exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X})\right] \left[\sum_{j=0}^{M} \frac{\partial \lambda_j}{\partial t} f_{\mu_j}(\boldsymbol{X})\right]
\end{aligned} \tag{6.5}$$

This equation and the definition of moments (equation (6.2)) connect the time derivative of the moments to time derivative of the Lagrange multipliers:

$$
\begin{aligned}
\frac{\partial \mu_i}{\partial t} &= \frac{\partial \sum_{\Omega} f_{\mu_i}(\boldsymbol{X}) P(\boldsymbol{X})}{\partial t} \\
&= \sum_{\Omega} f_{\mu_i}(\boldsymbol{X}) \frac{\partial P(\boldsymbol{X})}{\partial t} \\
&= -\sum_{\Omega} \left\{ f_{\mu_i}(\boldsymbol{X}) \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right] \left[\sum_{j=0}^{M} \frac{\partial \lambda_j}{\partial t} f_{\mu_j}(\boldsymbol{X}) \right] \right\} \\
&= -\sum_{\Omega} \left(\sum_{j=0}^{M} \left\{ f_{\mu_i}(\boldsymbol{X}) \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right] \frac{\partial \lambda_j}{\partial t} f_{\mu_j}(\boldsymbol{X}) \right\} \right) \\
&= -\sum_{\Omega} \left(\sum_{j=0}^{M} \left\{ f_{\mu_{i+j}}(\boldsymbol{X}) \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right] \frac{\partial \lambda_j}{\partial t} \right\} \right) \\
&= -\sum_{j=0}^{M} \left(\sum_{\Omega} \left\{ f_{\mu_{i+j}}(\boldsymbol{X}) \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right] \right\} \frac{\partial \lambda_j}{\partial t} \right)
\end{aligned}
\tag{6.6}
$$

By definition (equation (6.4)):

$$
\sum_{\Omega} \left\{ f_{\mu_{i+j}}(\boldsymbol{X}) \exp\left[-\sum_{j=0}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right] \right\} = \mu_{i+j} \tag{6.7}
$$

Thus, the moments' time derivative is transformed to:

$$
\frac{\partial \mu_i}{\partial t} = -\sum_{j=0}^{M} \left(\mu_{i+j} \frac{\partial \lambda_j}{\partial t} \right) \tag{6.8}
$$

The sum of the probability distribution across the whole state space is always 1 by definition. As a result, the zero$^{\text{th}}$ order moment is 1 ($\mu_0 = 1$) and its time derivative is zero ($\frac{\partial \mu_0}{\partial t} = 1$). This results in the zero$^{\text{th}}$ Lagrange multiplier λ_0 to be dependent on the rest of Lagrange multipliers

$$\begin{aligned}
0 &= \frac{\partial \mu_0}{\partial t} \\
&= -\sum_{j=0}^{M}\left(\mu_j \frac{\partial \lambda_j}{\partial t}\right) \\
&= -\mu_1 \frac{\partial \lambda_0}{\partial t} - \sum_{j=1}^{M}\left(\mu_j \frac{\partial \lambda_j}{\partial t}\right) \\
&= -\frac{\partial \lambda_0}{\partial t} - \sum_{j=1}^{M}\left(\mu_j \frac{\partial \lambda_j}{\partial t}\right)
\end{aligned} \tag{6.9}$$

Hence, the time derivative of the zero$^{\text{th}}$ Lagrange multiplier depends on the rest of the time derivatives as follows:

$$\frac{\partial \lambda_0}{\partial t} = -\sum_{j=1}^{M}\left(\mu_j \frac{\partial \lambda_j}{\partial t}\right) \tag{6.10}$$

Based on this relation, equation (6.8) can be modified as:

$$\begin{aligned}
\frac{\partial \mu_i}{\partial t} &= -\sum_{j=0}^{M}\left(\mu_{i+j} \frac{\partial \lambda_j}{\partial t}\right) \\
&= -\mu_i \frac{\partial \lambda_0}{\partial t} - \sum_{j=1}^{M}\left(\mu_{i+j} \frac{\partial \lambda_j}{\partial t}\right) \\
&= \mu_i \sum_{j=1}^{M}\left(\mu_j \frac{\partial \lambda_j}{\partial t}\right) - \sum_{j=1}^{M}\left(\mu_{i+j} \frac{\partial \lambda_j}{\partial t}\right) \\
&= \sum_{j=1}^{M}\left(-\mu_{i+j} \frac{\partial \lambda_j}{\partial t} + \mu_i \mu_j \frac{\partial \lambda_j}{\partial t}\right) \\
&= \sum_{j=1}^{M}\left(-\mu_{i+j} + \mu_i \mu_j\right) \frac{\partial \lambda_j}{\partial t}
\end{aligned} \tag{6.11}$$

We have previously proved (Chapter 5, section 5.3.1) that:

$$-\mu_{i+j} + \mu_i \mu_j = \frac{\partial \mu_i}{\partial \lambda_j} \tag{6.12}$$

Hence, the time derivative of one moment is given by (equations (6.11) & (6.12)):

$$\frac{\partial \mu_i}{\partial t} = \sum_{j=1}^{M} \frac{\partial \mu_i}{\partial \lambda_j} \frac{\partial \lambda_j}{\partial t} \tag{6.13}$$

or in matrix form:

$$\frac{\partial \boldsymbol{\mu}}{\partial t} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\lambda}} \frac{\partial \boldsymbol{\lambda}}{\partial t} = J \frac{\partial \boldsymbol{\lambda}}{\partial t} \tag{6.14}$$

J represents the Jacobian matrix of the system. Equation (6.14) also represents the chain rule of differentiation. There is now a way to connect the moment equations to equations that involve the Lagrange multipliers equations. The Lagrange multiplier equations can be derived by equation (6.14) and the moment equations (equation (6.1)):

$$\frac{\partial \boldsymbol{\lambda}}{\partial t} = J^{-1} \frac{\partial \boldsymbol{\mu}}{\partial t} = J(\lambda)^{-1} \left[A\boldsymbol{\mu}(\boldsymbol{\lambda}) + A'\boldsymbol{\mu}'(\boldsymbol{\lambda}) + \boldsymbol{\mu_c} \right] \tag{6.15}$$

where J^{-1} is the inverse of the Jacobian matrix. In the Appendix A.2, we discuss a computationally efficient way to calculate it. Even though the above equation also includes the probability moments, it only depends on the Lagrange multipliers since the moments are directly correlated to Lagrange multipliers (equation (6.4)).

All the Lagrange multipliers can be calculated through equation (6.15) except the zero$^{\text{th}}$ order one λ_0. As discussed earlier, λ_0 depends on the rest of the Lagrange multipliers through the relation (Chapter 5, section 5.3.1):

$$\lambda_0 = \log \left\{ \sum_{\Omega} \left[\exp \left(- \sum_{j=1}^{M} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right) \right] \right\} \tag{6.16}$$

We refer to equation (6.15) as the ”Lagrange multiplier equations (LMEs)”. To our knowledge, this is the first time that LMEs are being discussed and derived in the literature. Through this step, the problem of calculating the probability distribution for each point of the state space of stochastic networks has been transformed into a problem of calculating a finite set of Lagrange multipliers. The system is closed, it has the same number of equations and unknowns, and it is an initial value problem. An iterative numerical method can be used to solve this system.

6.2.2 Runge-Kutta

As mentioned, the LMEs problem is an initial value problem. A proposed way to solve the system is the Dormand-Prince RKF5(4)7M Runge-Kutta method [87]. RKF5(4)7M is an explicit numerical method with adaptive step size. Dormand and Prince formulated the method in order to create a stable approach for local extrapolations; this is the reason why it is chosen in this work as it will help to create more stable solution for time calculations by extrapolating the knowledge of past time solutions. A table that includes all the constants of the RKF5(4)7M method (often called ”Butcher tableau” [91] or simply ”tableau”) can be found in the Appendix A.3.

As derived in equation (6.15), the LMEs are:

$$\frac{\partial \boldsymbol{\lambda}}{\partial t} = J(\lambda)^{-1}\left[A\boldsymbol{\mu}(\boldsymbol{\lambda}) + A'\boldsymbol{\mu}'(\boldsymbol{\lambda}) + \boldsymbol{\mu_c}\right] = \boldsymbol{F}(\boldsymbol{\lambda}) \tag{6.17}$$

where $\boldsymbol{F}$ denotes the functional form of the LMEs.

The calculation of Lagrange multipliers, $\boldsymbol{\lambda}$, is desired for multiple time points. In order to calculate the multipliers at the time point $n+1$, the knowledge of the multipliers at time n is required. The Lagrange multipliers at time $n+1$, $\boldsymbol{\lambda_{n+1}}$, are calculated as:

$$\boldsymbol{\lambda_{n+1}} = \boldsymbol{\lambda_n} + \sum_{i=1}^{7} \hat{b}_i \boldsymbol{k_i} \tag{6.18}$$

where $\boldsymbol{\lambda_{n+1}}$ are the Lagrange multipliers at time n, $\hat{b}_i$ are constants of RKF5(4)7M and $\boldsymbol{k_i}$ are the Runge-Kutta coefficients:

$$\boldsymbol{k_i} = h\boldsymbol{F}\left(\boldsymbol{\lambda_n} + \sum_{j=1}^{i-1} a_{ij}\boldsymbol{k_j}\right) \tag{6.19}$$

h is the adaptive step size and a_{ij} are RKF5(4)7M constants.

In RKF5(4)7M, a lower-order solution, $\boldsymbol{\lambda}^*$, is used to evaluate the error of the solution $\boldsymbol{\lambda}$ and change accordingly the adaptive step size h. The lower-order solution at time $n+1$, $\boldsymbol{\lambda}^*_{\boldsymbol{n+1}}$, are calculated as:

$$\boldsymbol{\lambda}^*_{\boldsymbol{n+1}} = \boldsymbol{\lambda}^*_{\boldsymbol{n}} + \sum_{i=1}^{7} b_i \boldsymbol{k_i} \tag{6.20}$$

where b_i's are another set of RKF5(4)7M constants. All the constants (a_{ij},$\hat{b}_i$ & b_i) can be found in the Appendix A.3

More details about the algorithm of RKF5(4)7M used in this work can be found in section 7.2 of Chapter 7.

6.2.3 Initial Value from Moments to Lagrange Multipliers

The LMEs problem is an initial value problem; the values of the Lagrange multipliers for the initial state are required. However, Lagrange multipliers are numerical variables with no physical meanings and as a result their value is unknown. What is known is either the initial probability distribution or its probability moments. Hence, a way to calculate Lagrange multipliers from probability moments is required. It should be noted that there is not a universally acceptable way to back calculate probability distributions from moments and the problem is not trivial [92].

Let's assume that the moments, $\boldsymbol{\mu}\,(t=0)$ of the initial distribution, $P\,(t=0)$ are known. If only the initial distribution is known, its moments can be calculated from equation (6.2). The probability distribution moments are related to the Lagrange multipliers by equation (6.4). Hence:

$$
\begin{aligned}
\boldsymbol{\mu}\left(\boldsymbol{t=0}\right) &= \begin{bmatrix} \mu_1\left(t=0\right) \\ \mu_2\left(t=0\right) \\ \vdots \end{bmatrix} \\
&= \begin{bmatrix} \sum_{\Omega}\left\{f_{\mu_1(t=0)}(\boldsymbol{X})\exp\left[-\sum_{j=0}^{M}\lambda_j\left(t=0\right)f_{\mu_j(t=0)(t=0)}(\boldsymbol{X})\right]\right\} \\ \sum_{\Omega}\left\{f_{\mu_2(t=0)}(\boldsymbol{X})\exp\left[-\sum_{j=0}^{M}\lambda_j\left(t=0\right)f_{\mu_j(t=0)(t=0)}(\boldsymbol{X})\right]\right\} \\ \vdots \end{bmatrix} \\
&= \boldsymbol{G}\left[\boldsymbol{\lambda}\left(\boldsymbol{t=0}\right)\right]
\end{aligned}
\tag{6.21}
$$

$\boldsymbol{G}$ denotes the functional form of the matrix with dependences of the Lagrange multipliers. In order to calculate the Lagrange multipliers at the initial time from the probability moments, one has to solve a problem of the form: $\boldsymbol{G}\left(\boldsymbol{\lambda}\right) - \boldsymbol{\mu} = 0$. This is an infinite set and in order to be numerically solvable the user should decide the necessary number of lower-order moments (i.e. to specify the closure order [28]). When closure order is specified, the system has the same number of equations and unknowns (the Lagrange multipliers) and a root-finding method such as Newton-Rapshon can be used. The residual of the method is $\boldsymbol{R} = \boldsymbol{G}\left(\boldsymbol{\lambda}\right) - \boldsymbol{\mu}$ and the Jacobian matrix is $J_{i,j} = \frac{\partial F_i}{\partial \lambda_j} = \frac{\partial \mu_i}{\partial \lambda_j} = -\mu_{i+j} + \mu_i \mu_j$ [40]. With this approach all the Lagrange multipliers can be calculated other than the zero$^{\text{th}}$ one. The zero$^{\text{th}}$ one can be calculated from equation (6.16).

The knowledge of probability moments is used here to calculate the Lagrange multipliers, which then can be used to calculate the probability distribution based on equation (6.3). This approach allows to calculate probability distributions from its probability moments. The novelty of the method is the correlation of the probability distribution with Lagrange multipliers by maximizing the entropy of the system.

6.3 Results

In order to demonstrate the advantages of the proposed Lagrange multiplier equations (LMEs) approach, we employ four different networks: the bistable Schlögl model [46],

Table 6.1: The table shows the reaction networks with their kinetic constants for the bistable Schlögl, multistable Wilhelm, oscillatory Brusselator and viral infection networks. The kinetic constants values and initial condition for each system are also presented. For the third order reactions the kinetic constants are in $(molecules^2 \cdot s)^{-1}$ units, for the second order in $(molecules \cdot s)^{-1}$, for the first order in s^{-1} and for the zeroth order in $molecules \cdot s^{-1}$. Each system's initial distribution is a Gaussian distribution. The mean (first polynomial moment) is reported for each network.

Models	*Reactions*	*Kinetic constants*	*Initial Moments*
Schlögl	$3X \xrightarrow{k_1} 2X$	$k_1 = 1.5\ 10^{-3}$	$\{X\} = 38.01$
	$2X \xrightarrow{k_2} 3X$	$k_2 = 1.5\ 10^{-1}$	
	$X \xrightarrow{k_3} \emptyset$	$k_3 = 3.5$	
	$\emptyset \xrightarrow{k_4} X$	$k_4 = 22$	
Wilhelm	$Y \xrightarrow{k_1} 2X$	$k_1 = 35$	$\{X\} = 15.94$ $\{Y\} = 7.14$
	$2X \xrightarrow{k_2} X + Y$	$k_2 = 1$	
	$X + Y \xrightarrow{k_3} Y$	$k_3 = 1$	
	$X \xrightarrow{k_4} \emptyset$	$k_4 = 9.74$	
	$\emptyset \xrightarrow{k_5} X$	$k_5 = 30$	
Brusselator	$\emptyset \xrightarrow{k_1} X$	$k_1 = 10$	$\{X\} = 10.68$ $\{Y\} = 7.47$
	$2X + Y \xrightarrow{k_2} 3X$	$k_2 = 9\ 10^{-3}$	
	$X \xrightarrow{k_3} Y$	$k_3 = 3$	
	$X \xrightarrow{k_4} \emptyset$	$k_4 = 1$	
Viral Infection*	$D + P \xrightarrow{k_1} \emptyset$	$k_1 = 1$	$\{D\} = 15.41$ $\{P\} = 25.53$ $\{R\} = 49.56$
	$D \xrightarrow{k_2} R + D$	$k_2 = 3$	
	$R \xrightarrow{k_3} \emptyset$	$k_3 = 1$	
$*D = DNA$ $P = Protein$ $R = RNA$	$R \xrightarrow{k_4} D + R$	$k_4 = 10$	
	$R \xrightarrow{k_5} P + R$	$k_5 = 110$	
	$P \xrightarrow{k_6} \emptyset$	$k_6 = 200$	

the multistable Wilhelm's system [49], the oscillatory Brusselator [88, 89] and the viral infection [18, 90]. The reaction networks and kinetic constants for each model are reported in table 6.1. In order to assess the accuracy and computational efficiency of this method, the results are compared with SSA results using the same initial condition and kinetic constants.

For all the systems the initial condition is a Gaussian distribution, away from the steady state, generated with SSA. The mean (first polynomial moment [22]) of each initial distribution can be found in table 6.1. Using the algorithm in section 6.2.3, the initial distribution was transformed into an initial set of Lagrange multipliers. Based on the initial Lagrange multipliers and the Prince-Dormant algorithm outlined in section 7.2 of Chapter 7, the Lagrange multipliers for every time point were calculated until steady state was reached. Equations (6.3) and (6.4) were then used to calculate the probability distribution and polynomial moments for each time point based on the calculated Lagrange multipliers.

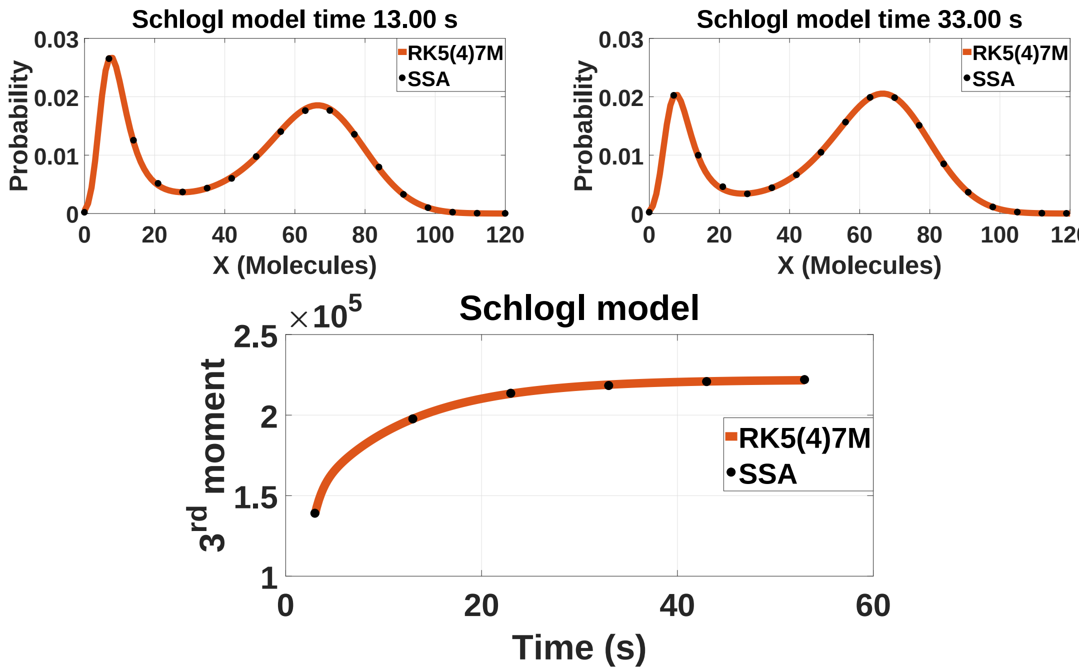


Figure 6.1: The probability distribution and the third order moment of the bistable Schlögl model. The first row presents the probability distribution of the single component for two different time points (13 s the left and 33 s after the initial condition the right one). The second row presents the time evolution of the thirds moment The solid lines are solutions calculated based on LMEs with tenth-order closure. The dots presents results from 500.000 SSA trajectories.

The Schlögl model is one of the most commonly used theoretical bistable systems and is the simplest single-component system that can exhibit bistability [47, 49]. For different kinetic constants, the model has either a unimodal or bimodal distribution. The kinetic constants selected in table 6.1 result in a bimodal network. Results for the bistable model are reported on figure 6.1. The probability distribution of the model is a bimodal distribution that changes form with time. The LMEs approach is able to capture the bistable behavior and its time evolution. Biological systems exhibiting bistability, both natural and synthetic, have gained an increasing interest recently [61, 63, 64], and the Schlögl model is a simple prototypical bistable model.

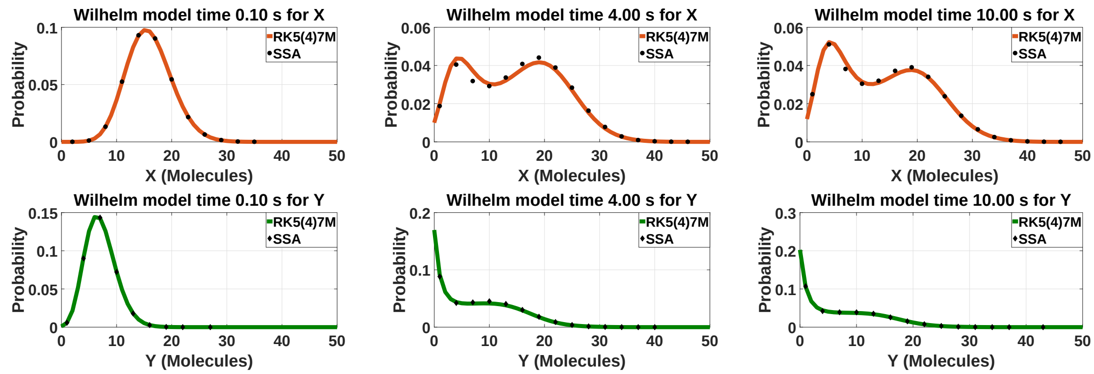


Figure 6.2: The probability distributions of the multistable Wilhelm's model. The first row presents results for component X and the second row for component Y. The first column presents the time close to the initial condition (0.1 s), the second a mid-time point (4 s) and the last column a time point close to steady-state (10 s). The solid lines are solutions calculated based on LMEs and the dots/diamonds SSA solutions. The LMEs are calculated with up to sixth-order moments. For SSA, 500.000 trajectories were used.

Another example of a theoretical system that can exhibit bistability is Wilhelm's network. The network is a non-linear two-component model and it was created by Wilhelm as the smallest known bistable chemical reaction system [49]. Depending on the kinetic constants, both components can exhibit bistability simultaneously which results in a multistable system overall. This is the case for the kinetic constants reported in table 6.1. The probability distribution for different time points for the two component of the network can be found in figure 6.2. The system starts with a unimodal distribution for both components. As the time progresses, both components have a different bimodal distribution resulting to a mutlistable network in total. Close to the steady state, both components still have bimodal distributions but different than before. The LMEs approach is able to capture all of these complex non-linear behaviors with accuracy similar to SSA.

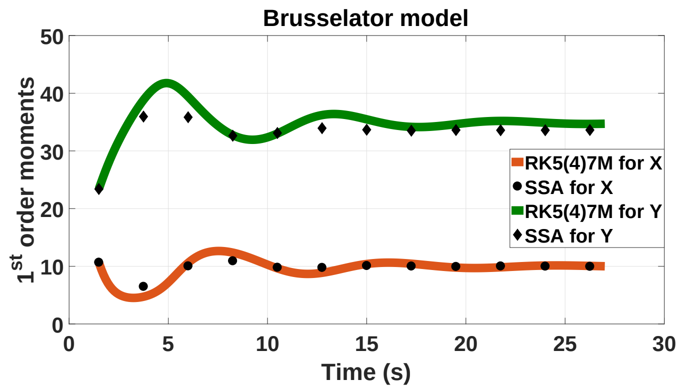


Figure 6.3: Time evolution of the first-order probability moments for the oscillatory Brusselator model. The solid-line upper green line presents the moment for Y component calculated based on LMEs. The solid-line lower orange line represents the moment for X. SSA results are also included from 500.000 trajectories; dots represent component X and diamonds component Y. The results are calculated with up to fourth-order moments.

Multistable non-linear behaviors are challenging to capture with the current stochastic tools, however oscillatory behaviors are even more challenging [93]. We have reported earlier how methods that can capture multistable behaviors fail to appropriately capture oscillatory ones (Chapter 5, [40]). The LMEs equations approach can be a solution to these numerical issues. In this work, we use Brusselator as a simple example of a theoretical oscillatory network [89]. We study the network in its oscillatory region as shown in figure 6.3. Both network's first-order moments oscillate with respect to time. The LMEs are able to capture this damped oscillatory behavior with acceptable accuracy.

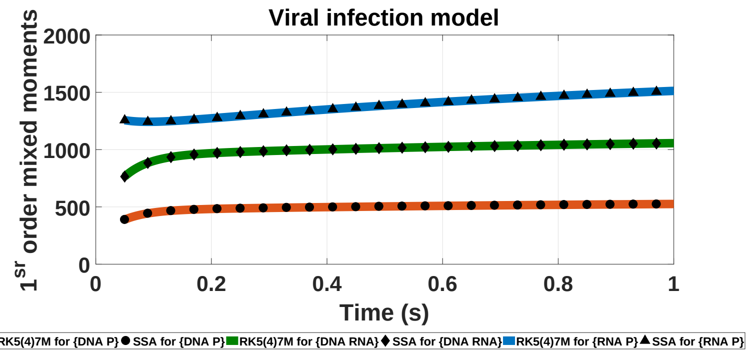


Figure 6.4: Time evolution of the first-order mixed probability moments for the Viral infection model. The solid-lines present results calculated based on LMEs and the doted points present SSA results. The results are calculated with up to second-order moment closure and 500.000 SSA trajectories. The upper blue line with triangles present the $\{DNA\ Protein\}$ mixed moment, the mid green line with diamonds corresponds to the $\{DNA\ RNA\}$ mixed moment and the lower orange line with dots the $\{RNA\ Protein\}$ mixed moment.

The LMEs approach is not only limited to theoretical systems with complex behavior but can also be applied to natural and synthetic biological networks. On such case of synthetic model is the viral infection model. The model was first created by Haseltine [18] as a general model of a cell infection by a virus. The model was then revised and simplified by Goutsias [90]. The network in table 6.1 reflects the revised version of Goutsias. Results for the mixed first-order moments are reported on figure 6.4. The LMEs approach is able to accurately capture the different dynamics of all the components and their combinations throughout the whole time evolution.

All the LMEs results obtained with RK5(4)M are compared with SSA results on table 6.2, in order to draw conclusions about the accuracy and the time efficiency of the method. Based on the Lagrange multipliers, the probability distribution and its polynomial moments were calculated for each time point until steady state is reached. The solution for each time point was compared to the one obtained from SSA simulations with 500,000 trajectories. The average difference among all time points between the two methods can be found on table 6.2. Results for both the probability distribution and the first two orders of polynomial moments are presented. For the system with more than one components (i.e. Wilhelm's, Brusselator and viral infection models), the moments of the same order are averaged. The probability difference is calculated with the Kullbalck - Leiber divergence [94], whereas the second norm is used to calculate the errors in the moments.

For all the systems, except the Brusselator, the LMEs solutions are as accurate as the SSA results. All the probability and moments errors are lower than 1%. The accuracy of the LMEs approach does not seem to be affected by the number of stability points of the system, since the two multistable systems (the Schlögl and the Wilhelm's models) have similar accuracy to the single stable one (the viral infection model). These systems also have different number of components and thus the number of components also does not affect the accuracy. Surprisingly, there is no significant difference between the first and

Table 6.2: The table presents the results of the LMEs solved with RK5(4)M compared to the ones obtained from SSA alongside the required computational time for each reaction network of table 6.1. Each SSA simulation used 500.000 trajectories. The first three columns show the average error between the two methods, which is calculated for each time point until steady state is reached and then averaged across all of them. For the probability distribution ($P(\boldsymbol{X})$), the Kullback - Leibler divergence [94] is used. For the first and second-order moments the second norm is used. The last three columns present the time that is required for each method. Each system has different initial and final time as well as different step size. In order to take into account these differences, we report the solving time per each method's time step in CPU seconds. The last column presents the ratio of SSA time required over the RK5(4)M time.

Network	*Average error (%)*			*Time per step (CPU s)*		
	$P(\boldsymbol{X})$	*1st-order moments*	*2nd-order moments*	*LMEs*	*SSA*	*Time Ratio*
Schlögl	0.05	0.22	0.22	0.02	419.2	20960
Wilhelm	0.61	0.56	0.71	0.15	732.75	4885
Brusselator	9.90	2.81	6.07	0.74	68.05	92
Viral Infection	0.43	0.01	0.02	8.97	64.96	7

second order moment errors. The approach seems to relatively maintain its accuracy with higher-order moments, which is not the case for the steady-state methods [40].

However, the method haves less accuracy for solving oscillatory systems (the Brusselator model). The Brusselator network has an increased error in all categories compared to the other networks. Even in this case though, the error in both probability and moments is less than 10%. In this case, unlike the other networks, there is a significant increase in the error of the second-order moments compared to the first-order ones.

The difference in the accuracy can be associated with the order of closure that is used to generate the LMEs (in this case 4, figure 6.3). It has been reported that the Brusselator requires more than fourth-order moments for accurate results [93]. The reason, that the low order is used, is not the method itself but the accuracy of the initial condition. The initial probability distribution was created with SSA and then, based on the algorithm described in section 6.2.3, it was translated into the initial Lagrange multipliers. The algorithm in section 6.2.3 created significant numerical accuracy issues for moment orders higher than 4 and we were not able to create initial condition for higher lower-order moments. The Brusselator results are not ideal, however given the inability of current methods to solve oscillatory systems, this is a step in the right direction. Additionally, the error is not inherent in the LMEs approach but rather in the algorithm that calculates the initial condition.

Aside from the accuracy of the approach, table 6.2 also includes the computational time that is required for RK5(4)M to solve the LMEs. In all the systems, the LMEs approach is significantly faster than SSA. Each model has different computational time steps that are used; in order ot take this difference into account, the table shows the time required for one time point per system. For each system, at least 10,000 time steps were used. Thus, the total amount of time that is required for LMEs approach is significantly lower than for SSA.

Among all systems, the LMEs approach requires less computational time to solve the two multistable systems (the Schlögl and the Wilhelm's models). On the other hand, SSA requires significantly more time for these two systems compared to the other systems. The LMEs approach is incomparably faster (at least 4000 times faster, table 6.2) and advantageous especially for multistable systems. As for oscillatory dynamics, they do not seem to affect the computational time of the LMEs approach.

The number of components, on the other hand, affects significantly the computational time of the LMEs approach. The Wilhelm's and Brusselator models (both two-component systems) are slower than the Schlögl model (one-component system) and faster that the viral infection model (three-component system). SSA's computational time is not affected by the number of components but only by the number of stable points. Even in the slowest case (the viral infection model, table 6.2), the LMEs approach is at least seven times faster than SSA with the same accuracy.

The proposed method is a numerical method to solve stochastic reaction networks. The LMEs include summations across all the possible values of the state space. The state space needs to be specified and be finite in order for the algorithm to perform numerical calculations. This is a known case with numerical methods [28, 95] and it can raise concerns about the accuracy of a method based on the state space size, especially in the case of systems with unknown state space. Some ways to calculate the appropriate state space of stochastic networks have been discussed in the literature [85, 95].

In order to evaluate the stability of the Lagrange multipliers approach with the state space size, we use the Schlögl model as an example. The model has one component and thus the connection between state space size and solution accuracy can be drawn easily; yet it has a complex enough bistable behavior. Figure 6.5 shows results of the LMEs approach for different state space sizes. The solutions are also compared with the SSA results. The accuracy of the method is not affected by the state space size. All different state space lengths provide the same accurate solution. Thus, the LMEs approach is also a robust alternative for systems with unknown state space.

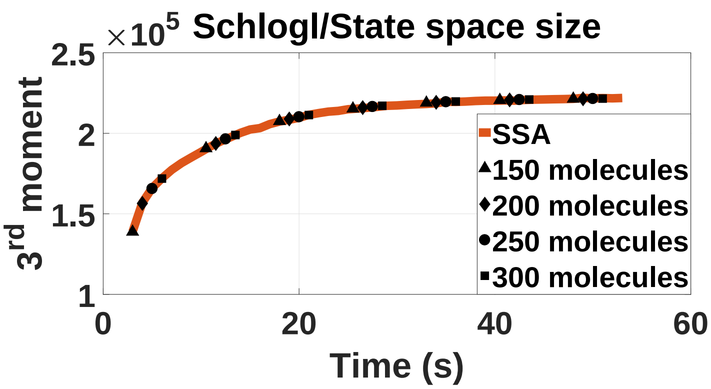


Figure 6.5: Time evolution of the third-order moment for the Schögl model. The plot is similar to the bottom plot of figure 6.1. Here, four different state space lengths are plotted. All the cases have zero as the lower bound of the state space. The coloring in this picture is reverse than the rest of them. The solid orange line is the SSA solution for 500.000 trajectories. The black dots are results with tenth-closure LMEs. The triangles represent the solution with 150 molecules as the upper bound of the state space, diamonds use 200 molecules limit, circles 250 molecules and the squares 300 molecules.

6.4 Conclusions

Stochasticity governs reaction networks away from the thermodynamic limit. The most common approach to solve stochastic reacting systems is through moment equations. However, moment equations are not a closed system and require important assumptions about the probability distribution of the system.

Here, we present an alternative method to solve stochastic reaction networks by utilizing the Lagrange multipliers. An approach to transform moment equations into a closed system of Lagrange multiplier equations (LMEs) is described. Sine, LMEs is a closed system, Runge-Kutta methods (e.g. Durmand-Prince RK5(4)7M) can be employed to solve the transient problem and calculate the Lagrange multipliers for different time points. With the knowledge of the Lagrange multipliers, one then can calculate the probability distribution and its moments at any given time.

This is a novel approach to bypass numerical and computational issues with Chemical Master Equation and moment equations. The only assumption used is that the entropy of the system is maximum at all times. The LMEs solutions are as accurate as alternative methods, such as SSA, with significant computational advantages. Four different non-linear reaction networks are employed to demonstrate the advantages of the new LMEs approach. The networks vary in component numbers (from one to three) and complexity (including multistability and oscillations). The number of stable solutions of the network does not affect the accuracy of the LMEs approach. Most notably, the method is computationally efficient for multistable systems, especially with the growing scientific interest in them [61,63,64]. It is also unaffected by the size of the system's state space; the method is stable and accurate for multiple state space sizes. Oscillatory behaviors can reduce its accuracy, which still remains within acceptable limits.

For the examples studied, the LMEs approach is always faster than SSA. The non-linear behavior of the network leaves the computational needs unchanged. On the other hand, the number of the network's components can affect the method's computational requirements.

Chapter 7

Algorithms

This chapter reports all the algorithms employed for Chapters 2, 3, 4 & 6.

7.1 Newton-Raphson Algorithm for Steady State Calculations with ZI Closure Scheme

The fundamental equations of ZI closure scheme have been reported in [1, 28, 39]. We have developed an updated version of the Newton-Raphson algorithm for the ZI closure scheme (Chapters 1 & 2). This section reports this update algorithm.

The steady state moment equations ($\mathbf{0} = A\boldsymbol{\mu} + A'\boldsymbol{\mu'} + \boldsymbol{\mu_c}$) depend only on Lagrange multipliers ($\boldsymbol{\lambda}$) and the kinetic constants of the system. The equation that connects moments and langange multipliers is: $\mu_i = \sum_\Omega f_{\mu_i}(\boldsymbol{X})P(\boldsymbol{X})$, where $f_{\mu_i}(\boldsymbol{X})$ represents the functional form of the moment i evaluated at $\boldsymbol{X}$ and Ω corresponds to the N-dimensional state space for all the possible values of $\boldsymbol{X}$. The system of equations is closed, i.e. it has the same number of unknowns (Lagrange multipliers) and equations (moment equations), and a root-finding method can be equipped in order to find the steady state Lagrange multipliers.

The residual of the method is calculated as $R = A\boldsymbol{\mu} + A'\boldsymbol{\mu'} + \boldsymbol{\mu_c}$ based on the moment equations. Matrices A, A' and the vector of constants $\boldsymbol{\mu_c}$ of the moment equations are generated as described in [22]. The Jacobian matrix is given by: $J_{i,j} = \frac{\partial \mu_i}{\partial \lambda_i} = -\mu_{i,j} + \mu_i \cdot \mu_j$, where $\mu_{i,j} = \sum_\Omega \left[f_{\mu_i}(\boldsymbol{X}) f_{\mu_j}(\boldsymbol{X}) P(\boldsymbol{X}) \right]$.

The proposed algorithm based on the Newton-Raphson method is:

(i) Set the necessary order of moments (Ψ). The value of Ψ can affect the accuracy of the solution since by setting a specific number of lower-order moments, we assume that the rest of them have little to offer in the solution.

(a) Calculate matrices A and A'. These matrices depend only on the kinetic constants of the system. The calculation of A and A' is detailed in [22].

(b) Introduce an initial guess for the Lagrange multipliers of the system $\boldsymbol{\lambda_\Psi} = [\lambda_1, \lambda_2, \ldots]$. The subscript Ψ denotes that the values of Lagrange multipliers refer to Ψ order of moments. Usually this guess is either a vector of zeros (which means that a uniform distribution is the initial guess) or the Lagrange multipliers generated from a system close to the one it is getting solved.

(c) Calculate the zero^th Lagrange multiplier, λ_0. λ_0 depends on the rest of them through the equation: $\lambda_0 = \log \left\{ \sum_\Omega \left[\exp \left(-\sum_{j>0} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right) \right] \right\} - 1$.

(d) Calculate the moment vectors $\boldsymbol{\mu_{\lambda_\Psi}}$ and $\boldsymbol{\mu'_{\lambda_\Psi}}$ based on equation $\mu_i = \sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp \left(-1 - \sum_j \lambda_j f_{\mu_j}(\boldsymbol{X}) \right)$ and the current value of the Lagrange multipliers $\boldsymbol{\lambda_\Psi}$. The subscript Ψ is employed to emphasize that the vectors have values based on the current guess of Lagrange multipliers.

(e) Calculate the value of the moment equations for the current guess of moments:

$$\Delta \boldsymbol{\mu_{\lambda_\Psi}} = A \boldsymbol{\mu_{\lambda_\Psi}} + A' \boldsymbol{\mu'_{\lambda_\Psi}} + \boldsymbol{\mu_c} \tag{7.1}$$

(f) Calculate the Euclidean norm of vector $\Delta \boldsymbol{\mu_{\lambda_\Psi}}$:

$$\epsilon = \sqrt{\Delta\boldsymbol{\mu}^T_{\boldsymbol{\lambda}_\Psi}\Delta\boldsymbol{\mu}_{\boldsymbol{\lambda}_\Psi}}. \tag{7.2}$$

(g) If ϵ is less than the desired accuracy threshold proceed to (ih). Otherwise:

- Calculate the Jacobian of the system:

$$J = \frac{\partial\Delta\boldsymbol{\mu}_{\boldsymbol{\lambda}_\Psi}}{\partial\boldsymbol{\lambda}_\Psi} = A\frac{\partial\boldsymbol{\mu}_{\boldsymbol{\lambda}_\Psi}}{\partial\boldsymbol{\lambda}_\Psi} + A'\frac{\partial\boldsymbol{\mu}'_{\boldsymbol{\lambda}_\Psi}}{\partial\boldsymbol{\lambda}_\Psi}. \tag{7.3}$$

- Based on Newton-Raphson method, calculate a first order Taylor expansion:

$$\Delta\boldsymbol{\mu}_{\boldsymbol{\lambda}_\Psi} = J\Delta\boldsymbol{\lambda}_\Psi. \tag{7.4}$$

- Create an approximate step $\Delta\boldsymbol{\lambda}_\Psi$ for the Lagrange multipliers vector $\boldsymbol{\lambda}_\Psi$:

$$\Delta\boldsymbol{\lambda}_\Psi = J^{-1}\Delta\boldsymbol{\mu}_{\boldsymbol{\lambda}_\Psi}. \tag{7.5}$$

  where J^{-1} is the inverse of the Jacobian matrix J. Appendix A.2 provides some numerical insights for a computationally efficient way to calculate J^{-1}.

- Calculate the new set of Lagrange multipliers: $\boldsymbol{\lambda}_\Psi = \boldsymbol{\lambda}_\Psi + \Delta\boldsymbol{\lambda}_\Psi$.
- Return to (id).

(h) The final solution for Ψ order of moments is $\boldsymbol{\lambda}_\Psi$ and the probability distribution $\boldsymbol{P}_\Psi$. $\boldsymbol{P}_\Psi$ is calculated based on $\boldsymbol{\lambda}_\Psi$ and equation $P_\Psi(y) =$
$= \exp\left[-1 - \sum_j \lambda_j f_{\mu_j}(\boldsymbol{X})\right]$.

The final solution should be independent of Ψ. The following steps of the algorithm ensure that.

(ii) Increase order of moments by 1. Return to ia. At step ib, the initial guess of the Lagrange multipliers ($\boldsymbol{\lambda}_{\Psi+1}$) is the old calculated values accompanied with zeros in order to maintain proper vector size ($\boldsymbol{\lambda}_{\boldsymbol{\Psi}+\mathbf{1}} = [\boldsymbol{\lambda}_{\boldsymbol{\Psi}}, 0, 0, \ldots, 0]$). Calculate the final Lagrange multipliers for the new order of moments ($\boldsymbol{\lambda}_{\boldsymbol{\Psi}+\mathbf{1}}$) and each associate probability distribution $\boldsymbol{P}_{\boldsymbol{\Psi}+\mathbf{1}}$.

(iii) Calculate the difference between the probability distributions for different order moments

$$\Delta \boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{\Psi}+\mathbf{1}} - \boldsymbol{P}_{\boldsymbol{\Psi}} \tag{7.6}$$

(iv) Calculate the Euclidean norm of vector $\Delta \boldsymbol{P}$

$$\epsilon_P = \sqrt{\Delta \boldsymbol{P}^T \Delta \boldsymbol{P}}. \tag{7.7}$$

(v) If ϵ_P is less than the wanted accuracy threshold, the solution is independent of the moment order and the user can proceed to step vi. Otherwise, return to ii and repeat calculations.

(vi) The final order of moments is $\Psi_{final} = \Psi + 1$ (or whichever order gives ϵ_P less than the desired accuracy). The final solution is $\boldsymbol{\lambda}_{\boldsymbol{\Psi}_{final}}$ and the probability distribution is $\boldsymbol{P}_{\boldsymbol{\Psi}_{final}}$.

7.2 Runge-Kutta Algorithm

Chapter 6 discusses how through the Lagrange multiplier equations (LMEs) one can calculate the time transient behavior of stochastic systems. This approach is innovative and a major improvement compared to the steady state moment equation algorithms, *e.g.* ZI closure scheme. The proposed algorithm to solve LMEs based on the Dormand-Prince RKF5(4)M method is:

(i) Calculate vector $\boldsymbol{\mu_c}$ and matrices A and A' as in [22]. These matrices depend only on the kinetic constants of the system and they are constant for all time points.

(ii) Introduce the time step size h. The time point of this iteration is $t_{n+1} = t_n + h$, where t_n is the time point of the previous iteration.

(iii) Increase the iteration counter by 1. For the rest of the algorithm the iteration step will be denoted by $n + 1$. At the first iteration, $n = 0$.

(iv) Introduce the previous step (iteration n) Lagrange multipliers $\boldsymbol{\lambda_n}$. At the first iteration, introduce the initial condition $\boldsymbol{\lambda_0}$. The subscripts here refers to the time step and are applied on the whole Lagrange multiplier vector. They should not be confused with subscripts denoting specific Lagrange multiplier orders of the previous section.

(v) Calculate the $\boldsymbol{k_i}$'s coefficients from equation $\boldsymbol{k_i} = h\boldsymbol{F}\left(\boldsymbol{\lambda_n} + \sum_{j=1}^{i-1} a_{ij}\boldsymbol{k_j}\right)$ and the tableau (Appendix A.3). Each coefficient $\boldsymbol{k_i}$ depends on the functional form $\boldsymbol{F}$ of the Lagrange multiplier equations ($\boldsymbol{F}(\boldsymbol{\lambda}) = J(\lambda)^{-1}\left[A\boldsymbol{\mu}(\boldsymbol{\lambda}) + A'\boldsymbol{\mu'}(\boldsymbol{\lambda}) + \boldsymbol{\mu_c}\right]$) and on different set of Lagrange multipliers ($\boldsymbol{\lambda_n} + \sum_{j=1}^{i-1} a_{ij}\boldsymbol{k_j}$). In order to account for the change of the Lagrange multipliers, there is a need for a subroutine to calculate $\boldsymbol{k_i}$.

(vi) This is a proposed subroutine for $\boldsymbol{k_i}$

   (a) Calculate the augmented Lagrange multipliers for the specific $\boldsymbol{k_i}$, $\boldsymbol{\lambda_{n,k_i}} = \boldsymbol{\lambda_n} + \sum_{j=1}^{i-1} a_{ij}\boldsymbol{k_j}$. The values of a_{ij} can be found in the tableau in Appendix A.3.

   (b) Calculate the lower ($\boldsymbol{\mu}(\boldsymbol{\lambda_{n,k_i}})$) and higher-order moments ($\boldsymbol{\mu'}(\boldsymbol{\lambda_{n,k_i}})$) for the augmented Lagrange multipliers based on equation $\mu_i =$
   $= \sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp\left(-1 - \sum_j \lambda_j f_{\mu_j}(\boldsymbol{X})\right)$.

(c) Calculate the inverse of the Jacobian, $J(\lambda_{n,k_i})^{-1}$ for the augmented multipliers as described in Appendix A.2.

(d) Calculate the functional form $\boldsymbol{F}(\boldsymbol{\lambda_{n,k_i}})$ of the Lagrange multiplier equations ($\boldsymbol{F}(\boldsymbol{\lambda}) = J(\lambda)^{-1}\left[A\boldsymbol{\mu}(\boldsymbol{\lambda}) + A'\boldsymbol{\mu}'(\boldsymbol{\lambda}) + \boldsymbol{\mu_c}\right]$).

(e) Calculate $\boldsymbol{k_i}$ based on equations $\boldsymbol{k_i} = h\boldsymbol{F}(\boldsymbol{\lambda_{n,k_i}})$.

(vii) Repeat step vi for all seven $\boldsymbol{k_i}$ coefficients, $i = 1, \ldots, 7$. It's coefficient $\boldsymbol{k_i}$ depends on all the previous coefficients, thus the calculation of each one needs to be in sequence and include the subroutine. For example, the calculations of step vi should first performed for $i = 1$, then based on $\boldsymbol{k_1}$ to be performed for $i = 2$, then based on $\boldsymbol{k_2}$ to be performed for $i = 3$ etc.

(viii) Calculate the Lagrange multipliers for the current iteration $n + 1$, $\boldsymbol{\lambda_{n+1}}$, based on equation $\boldsymbol{\lambda_{n+1}} = \boldsymbol{\lambda_n} + \sum_{i=1}^{7} \hat{b}_i\boldsymbol{k_i}$ and the tableau (Appendix A.3).

(ix) Calculate the lower-order solution for the current iteration $n + 1$, $\boldsymbol{\lambda^*_{n+1}}$, based on equation $\boldsymbol{\lambda^*_{n+1}} = \boldsymbol{\lambda_n} + \sum_{i=1}^{7} b_i\boldsymbol{k_i}$ and the tableau (Appendix A.3).

(x) Check the accuracy of the solution by calculating the error ϵ:

$$\epsilon = \left|\frac{\boldsymbol{\lambda_{n+1}} - \boldsymbol{\lambda^*_{n+1}}}{\boldsymbol{\lambda_{n+1}}}\right|_2 \tag{7.8}$$

(xi) If the error ϵ is smaller than the wanted tolerance accept the solution and proceed to the next step. Otherwise, set a new time step size $h = h \cdot s$ and go back to step v and re-perform the calculation with the new step size. s is used to adjust the step size based on the accuracy of the solution and is given by:

$$s = \left(\frac{tolerance \cdot h}{2 \cdot |\boldsymbol{\lambda^*_{n+1}} - \boldsymbol{\lambda_{n+1}}|_2}\right)^{\frac{1}{4}} \tag{7.9}$$

(xii) Accept the solution of Lagrange multipliers for the current iteration $(n + 1)$ and current time point (t_{n+1}) as calculated at step viii.

(xiii) Set the new step size to $h = h \cdot s$ and proceed to step iii to calculate the solution for the rest of the time points until the final is reached.

(xiv) For any given time t, the probability distribution and its moments can be calculated based on the Lagrange multipliers $(\boldsymbol{\lambda(t)})$ and equations $P(\boldsymbol{X}) =$
$= \exp\left[-1 - \sum_j \lambda_j f_{\mu_j}(\boldsymbol{X})\right]$ & $\mu_i = \sum_\Omega f_{\mu_i}(\boldsymbol{X}) \exp\left(-1 - \sum_j \lambda_j f_{\mu_j}(\boldsymbol{X})\right)$.

Chapter 8

Concluding Remarks and Future Directions

In chemical engineering processes involving reacting systems, the canonical choice involves a continuous-deterministic modeling formalism, thereupon reaction kinetics of homogeneous systems are modeled with ordinary differential equations. An important hypothesis underlying the choice of a continuous-deterministic model is that the studied system is at the thermodynamic limit. At this limit the size of the system and the number of molecules of chemical species in the system are large enough for the concentration to be accurately represented with a continuously changing variable and for the ensemble of individual reaction events to average out to the deterministic reaction rate.

More recently stochastic reaction models have been the focus of numerous studies because they describe important aspects of biological processes more accurately than deterministic reaction kinetic models. In biomolecular systems, the number of molecules of reacting species is often very small. Stochasticity then impacts the observed cellular phenotypes, requiring modeling formalisms that consider the inherent variability in biological systems. Modeling these translation events with ordinary differential equations may be distinctly false.

The Chemical Master Equation (CME) is the most detailed mathematical model to describe such stochastic behavior, but its use is numerically intractable when it is applied to anything but the simplest of systems. While a plethora of kinetic Monte-Carlo algorithms (*e.g.* SSA) have been developed to numerically sample the master probability distribution, these algorithms carry significant computational cost.

An alternative approach is to study the moments of the probability distribution. There are numerous successful algorithms to generate moment equations. For non-linear systems, lower-order moments depend on higher-order ones. As a result, moment equations are challenging to solve without a closure scheme that relates higher and lower-order moments. This work suggests to close the moment equation scheme by relating all the moments to a set of Lagrange multipliers through maximizing the system's entropy. The resulting algorithm, Zero-Information (ZI) closure scheme, is a computationally efficient way to calculate the stationary probability distribution of non-linear biochemical reaction networks. Chapter 2 discusses in detail the use of factorial moments and the ZI closure scheme for solving stochastic reaction networks. The Schlögl model, a simple single-component system, is used to demonstrate all the mathematical aspects of the method.

There are numerous challenges we face when using ZI closure scheme. An important one is that the scheme has been tested only on systems with a small number of components (up to four); that is, networks with up to four chemical degrees of freedom. This is a practical, numerical implementation challenge, a result of the exploding phase space size for multidimensional problems. This is a numerical issue integrated into the method. In a future direction, the method should be applied on systems with more components using a more robust approach about truncating the state space. There are areas of the state space that provide minimal information to the overall system and there is a need to accurately identify them.

Another important limitation is that there is currently no way to know *a priori* how many moments suffice for a numerically accurate reconstruction of the probability distribution. Furthermore, the reconstruction itself becomes a progressively more challenging task as the number of moments increases. Numerically we have tackled the issue by comparing solutions of different moment orders until the difference is computational insignificant. A mathematically based approach would more ideal and save significant computational time.

Nevertheless, with ZI closure scheme there is now a tool to compute steady state probability distributions, without having to resort to stochastic simulations in time. This is an important advantage of the method compared to other stochastic simulation methods. Furthermore, with ZI closure scheme the tools become available to explore the stability of reaction network steady states and the sensitivity of the behavior to changes in system parameters, paving another avenue for investigating stochastic reaction networks.

In order to improve the accessibility to the scheme, Chapter 3 presents ZICS, an open access standalone Windows application, that employees it. With minimal computational cost, ZICS accurately calcualtes the stationary probability distribution of multistable multidimensional networks. ZICS also offers the ability to generate the factorial moment equations of a given network. ZICS is designed with the Matlab App deisgner. For non-Windows machines, the application can be used through Matlab, version 2016b and later.

As mentioned, the thermodynamic limit is an important concept for chemical reacting systems. Using the aforementioned tools, in Chapter 4 we explore how the system's size affects its probability distribution; how the distribution changes as the system transitions from the stochastic area to the thermodynamic limit. The Schlögl reaction network is used as a case study model. The network is a bistable system at the thermodynamic limit and exhibits bimodality at the stochastic region. The results cover a wide variety of system sizes, previously unattainable with traditional methods (e.g. SSA, analytical solution). This is possible now with the employment of ZI closure scheme. As the size increases, the exhibited bimodality collapses and only one of the attractors becomes dominant with probability 1.

It is revealed that the range of kinetic constant values where bimodality survives in larger systems diminishes in size. Presumably, there is a single value of the kinetic constant where bimodality survives at the thermodynamic limit. It is also found that the dominant eigenvalue of the stochastic system asymptotically reaches the eigenvalue of the corresponding deterministic solution as the volume increases. This implies that at the thermodynamic limit the probability distribution not only replicates one of the deterministic solutions but also fully regains the dynamic behavior of it as well. Thus at the thermodynamic limit, the bimodality collapses and the remaining unimodal distribution fully represents the behavior of one of the deterministic solutions.

The ZI closure scheme postulates the maximum entropy stationary probability distribution in order to numerically close the probability distribution moment equations. In the future, an investigation is needed to decide whether this criterion holds generally for non-equilibrium steady state systems. The scheme can be used to study the volume dependence of systems with more complex dynamics (*e.g.* limit cycles) and more components.

ZI closure scheme is only able to solve for stationary distribution. Chapter 5 & 6 emphasize on the development of methods for time dynamic calculations. Linearization methods are numerical approximations, useful for calculating time transient behaviors based on the knowledge of a system's steady state. In Chapter 5, we propose a linearization technique for solving the probability moment equation of stochastic reaction networks.

The approach linearizes moment equations through a two-term Taylor expansion. The time behavior of probability moments is then calculated through the known analytical solution of the linearized moment equations. The solution requires the eigenvalues and eigenvectors of the Jacobian matrix around the steady state solution. Thus, the calculation of the stationary Jacobian for the stationary probability distribution is a necessary intermediate step, for which we implement with the ZI closure scheme.

Three different stochastic reaction networks are studied to demonstrate the method: a network with two components and a non-linear behavior (Michaelis-Menten), a network with a single component and bistable behavior (the Schlögl model), and a two-component network with a more complicated bistable behavior (the Wilhelm model). In all three cases, the calculated moment behavior matches those calculated with SSA.

The approach is a linear approximation around the steady state. As such, the accuracy of the method depends on how close the initial conditions are from the steady state solution. The closer the initial guess is to the steady state solution, the more accurately the method will predict the time evolution of the moments. In the future, it will be beneficial to explore how the difference of the intial condition from the steady state affects the accuracy of the linearizion technique.

In Chapter 5, only the first two terms of the Taylor expansion are employed for the derivation of the linearized equation. As the higher order terms of the Taylor expansion become significant, the accuracy of the method may decrease. Systems with limit cycles are a prime example where higher order Taylor terms are important. In such cases, more terms of the Taylor expansion may be necessary. A new study of the effect of including more terms on the computational cost and accuracy of the method may be contacted.

In order to bypass all the numerical approximations of moment equations and the linearization technique, in Chapter 6 we propose a more robust approach for solving for the time probability distribution of stochastic networks. We present an alternative method to solve stochastic reaction networks by utilizing the Lagrange multipliers. An approach to transform moment equations into a closed system of Lagrange multiplier equations (LMEs) is described. Sine, LMEs is a closed system, Runge-Kutta methods (e.g. Durmand-Prince RK5(4)7M) can be employed to solve the transient problem and calculate the Lagrange multipliers for different time points. With the knowledge of the Lagrange multipliers, one then can calculate the probability distribution and its moments at any given time.

This is an innovative approach that only assumes maximum system's entropy at all times. The LMEs solutions are as accurate as alternative methods, such as SSA, with significant computational advantages. Four different non-linear reaction networks are employed to demonstrate the advantages of the new LMEs approach. The networks vary in component numbers (from one to three) and complexity (including multistability and oscillations). The number of stable solutions of the network does not affect the accuracy of the LME approach. Most notably, the method is computationally efficient for multi-stable systems. It is also unaffected by the size of the system's state space; the method is stable and accurate for multiple state space sizes. Oscillatory behaviors can reduce its accuracy, which still remains within acceptable limits. A future direction of the work can be to employ the method for networks with a large number of components.

All the algorithms of this work can be found in Chapter 7. The derivations in this work assume elementary reaction rates. There is a plethora of biological systems that utilize more complex rate dynamics. In such cases, the form of the moment equations and their generation algorithm should be altered. As a result, all the algorithms of this work need to reevaluated. A future study should be contacted in order to investigate this more complex behavior.

Bibliography

[1] P H Constantino, M Vlysidis, P Smadbeck, and Y N Kaznessis. Modeling stochasticity in biochemical reaction networks. *Journal of Physics D: Applied Physics*, 49(9):093001, mar 2016.

[2] Doraiswami Ramkrishna and Neal R. Amundson. Mathematics in chemical engineering: A 50 year introspection. *AIChE Journal*, 50(1):7–23, jan 2004.

[3] H.Scott Folger. *Elements of Chemical Reaction Engineering*. Prentice Hall, fourth edition, 2016.

[4] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annual review of physical chemistry*, 58:35–55, jan 2007.

[5] Donald A. McQuarrie. Stochastic approach to chemical kinetics. *Journal of Applied Probability*, 4(03):413–478, dec 1967.

[6] Irwin Oppenheim and Kurt E. Shuler. Master Equations and Markov Processes. *Physical Review*, 138(4B):B1007–B1011, may 1965.

[7] NG Van Kampen. *Stochastic processes in physics and chemistry*. Elsevier, fifth edition, 2004.

[8] Daniel T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(1-3):404–425, sep 1992.

[9] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, dec 1976.

[10] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, dec 1977.

[11] Michael A. Gibson and Jehoshua Bruck. Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *The Journal of Physical Chemistry A*, 104(9):1876–1889, mar 2000.

[12] Yang Cao, Hong Li, and Linda Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *The Journal of Chemical Physics*, 121(9):4059–4067, sep 2004.

[13] Larry Lok and Roger Brent. Automatic generation of cellular reaction networks with Moleculizer 1.0. *Nature Biotechnology*, 23(1):131–136, jan 2005.

[14] James M. McCollum, Gregory D. Peterson, Chris D. Cox, Michael L. Simpson, and Nagiza F. Samatova. The sorting direct method for stochastic simulation of biochemical systems with varying reaction execution behavior. *Computational Biology and Chemistry*, 30(1):39–49, feb 2006.

[15] Yiannis N Kaznessis. Multiscale models for synthetic biology. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2009:6408–11, jan 2009.

[16] Christopher V. Rao and Adam P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *The Journal of Chemical Physics*, 118(11):4999–5010, mar 2003.

[17] Jacek Puchałka and Andrzej M Kierzek. Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks. *Biophysical journal*, 86(3):1357–72, mar 2004.

[18] Eric L. Haseltine and James B. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *The Journal of Chemical Physics*, 117(15):6959–6969, oct 2002.

[19] Howard Salis, Vassilios Sotiropoulos, and YiannisN Kaznessis. Multiscale Hy3S: Hybrid stochastic simulation for supercomputers. *BMC Bioinformatics*, 7(1):93, feb 2006.

[20] Anthony D. Hill, Jonathan R. Tomshine, Emma M. B. Weeding, Vassilios Sotiropoulos, and Yiannis N. Kaznessis. SynBioSS: the synthetic biology modeling suite. *Bioinformatics*, 24(21):2551–2553, nov 2008.

[21] V. Sotiropoulos, M.-N. Contou-Carrere, P. Daoutidis, and Y.N. Kaznessis. Model Reduction of Multiscale Chemical Langevin Equations: A Numerical Case

Study. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 6(3):470–482, jul 2009.

[22] P Smadbeck and Y N Kaznessis. Efficient Moment Matrix Generation for Arbitrary Chemical Networks. *Chemical engineering science*, 84:612–618, dec 2012.

[23] Vassilios Sotiropoulos and Yiannis N Kaznessis. Analytical Derivation of Moment Equations in Stochastic Chemical Kinetics. *Chemical engineering science*, 66(3):268–277, feb 2011.

[24] Morris H. DeGroot and Mark J. Schervish. *Probability and Statistics*. Pearson, custom edi edition, 2012.

[25] C.S. Gillespie. Moment-closure approximations for mass-action models. *IET Systems Biology*, 3(1):52–58, jan 2009.

[26] J. E. Moyal. Stochastic Processes and Statistical Physics, 1949.

[27] D. A. McQuarrie, C. J. Jachimowski, and M. E. Russell. Kinetics of Small Systems. II. *The Journal of Chemical Physics*, 40(10):2914, 1964.

[28] Patrick Smadbeck and Yiannis N Kaznessis. A closure scheme for chemical master equations. *Proceedings of the National Academy of Sciences of the United States of America*, 110(35):14261–5, aug 2013.

[29] Abhyudai Singh and João P. Hespanha. Approximate Moment Dynamics for Chemically Reacting Systems. *IEEE Transactions on Automatic Control*, 56(2):414–418, feb 2011.

[30] H. Ted Davis and Kendall T. Thomson. *Linear Algebra and Linear Operators in Engineering*. Academic Press Inc, 2000.

[31] Angelique Ale, Paul Kirk, and Michael P H Stumpf. A general moment expansion method for stochastic kinetic models. *Journal of Chemical Physics*, 138(17), 2013.

[32] Peter Milner, Colin S. Gillespie, and Darren J. Wilkinson. Moment closure based parameter inference of stochastic kinetic models. *Statistics and Computing*, 23(2):287–295, 2013.

[33] Eszter Lakatos, Angelique Ale, Paul D. W. Kirk, and Michael P. H. Stumpf. Multivariate moment closure techniques for stochastic kinetic models. *The Journal of Chemical Physics*, 143(9):094107, sep 2015.

[34] Christoph Zechner, Jakob Ruess, Peter Krenn, Serge Pelet, Matthias Peter, John Lygeros, and Heinz Koeppl. Moment-based inference predicts bimodality in transient gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 109(21):8340–5, may 2012.

[35] A. Singh and J.P. Hespanha. Moment closure techniques for stochastic models in population biology. In *2006 American Control Conference*, page 6 pp. IEEE, 2006.

[36] Ramon Grima. A study of the accuracy of moment-closure approximations for stochastic chemical kinetics. *The Journal of chemical physics*, 136(15):154105, apr 2012.

[37] Philipp Thomas and Ramon Grima. Approximate probability distributions of the master equation. *Physical Review E*, 92(1):012120, jul 2015.

[38] Isthrinayagy Krishnarajah, Alex Cook, Glenn Marion, and Gavin Gibson. Novel moment closure approximations in stochastic epidemics. *Bulletin of Mathematical Biology*, 67:855–873, 2005.

[39] M. Vlysidis, P. H. Constantino, and Y. N. Kaznessis. ZI-Closure Scheme: A Method to Solve and Study Stochastic Reaction Networks. In *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology*, pages 159–174. Springer International Publishing, Cham, 2017.

[40] Michail Vlysidis and Yiannis N. Kaznessis. A linearization method for probability moment equations. *Computers & Chemical Engineering*, 112:1–5, apr 2018.

[41] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, jul 1948.

[42] J.N. Kapur. *Maximum-Entropy Models in Science and Engineering*. New age international publishers, 1989.

[43] James Gleick. *The information : a history, a theory, a flood*. Vintage, first edition, 2011.

[44] E. T. Jaynes. Information Theory and Statistical Mechanics. *Physical Review*, 106(4):620–630, may 1957.

[45] E. T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, first edition, 2003.

[46] F Schlogl. On Thermodynamics Near a Steady State. *Z. Physik*, 458:446–458, 1971.

[47] Daniel T. Gillespie. *Markov Processes, an introduction for physical scientists*. Academic Press Inc, 1992.

[48] Patrick Smadbeck and Yiannis N. Kaznessis. On a theory of stability for nonlinear stochastic chemical reaction networks. *The Journal of Chemical Physics*, 142(18):184101, may 2015.

[49] Thomas Wilhelm. The smallest chemical reaction system with bistability. *BMC Systems Biology*, 3(1):90, sep 2009.

[50] Bruce Alberts, Dennis Bray, Karen Hopkin, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Essentail Cell Biology*. Garland Science, New York, second edition, 2009.

[51] William W Chen, Mario Niepel, and Peter K Sorger. Classic and contemporary approaches to modeling biochemical reactions. *Genes & development*, 24(17):1861–75, sep 2010.

[52] Eberhard O. Voit, Harald A. Martens, and Stig W. Omholt. 150 Years of the Mass Action Law. *PLoS Computational Biology*, 11(1):e1004012, jan 2015.

[53] Thomas G. Kurtz. Limit theorems and diffusion approximations for density dependent Markov chains. In *Stochastic Systems: Modeling, Identification and Optimization, I Mathematical Programming Studies, Volume 5*, volume 5, pages 67–78. Springer, Berlin, Heidelberg, 1976.

[54] Thomas G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and their Applications*, 6(3):223–240, feb 1978.

[55] M Malek Mansour, C Van Den Broeck, G Nicolis, and J.W Turner. Asymptotic properties of markovian master equations. *Annals of Physics*, 131(2):283–313, feb 1981.

[56] H. K. Leung. Metastable states in a nonlinear stochastic model. *Physical Review A*, 30(5):2609–2612, nov 1984.

[57] Melissa Vellela and Hong Qian. A Quasistationary Analysis of a Stochastic Chemical Reaction: Keizer's Paradox. *Bulletin of Mathematical Biology*, 69(5):1727–1746, jun 2007.

[58] G. Nicolis and J. W. Turner. EFFECT OF FLUCTUATIONS ON BIFURCATION PHENOMENA. *Annals of the New York Academy of Sciences*, 316(1):251–262, feb 1979.

[59] I. Matheson, D. F. Walls, and C. W. Gardiner. Stochastic models of firstorder nonequilibrium phase transitions in chemical reactions. *Journal of Statistical Physics*, 12(1):21–34, jan 1975.

[60] Mark Ptashne. *A Genetic Switch, Phage Lambda Revisited*. Sloan-Kettering Memorial Cancer Center, New York, third edition, 2004.

[61] Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339–342, jan 2000.

[62] Yu-Chiun Wang and Edwin L. Ferguson. Spatial bistability of Dpp–receptor interactions during Drosophila dorsal–ventral patterning. *Nature*, 434(7030):229–234, mar 2005.

[63] David Angeli, James E Ferrell, Eduardo D Sontag, and Eduardo D. Sontag. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences of the United States of America*, 101(7):1822–7, feb 2004.

[64] Maxim N Artyomov, Jayajit Das, Mehran Kardar, and Arup K Chakraborty. Purely stochastic binary decisions in cell signaling models without underlying deterministic bistabilities. *Proceedings of the National Academy of Sciences of the United States of America*, 104(48):18958–63, nov 2007.

[65] Jeff Hasty, David McMillen, Farren Isaacs, and James J. Collins. Computational studies of gene regulatory networks: <i>in numero</i> molecular biology. *Nature Reviews Genetics*, 2(4):268–279, apr 2001.

[66] Hao Ge and Hong Qian. Thermodynamic Limit of a Nonequilibrium Steady State: Maxwell-Type Construction for a Bistable Biochemical System. *Physical Review Letters*, 103(14):148103, oct 2009.

[67] Hao Ge and Hong Qian. Non-equilibrium phase transition in mesoscopic biochemical systems: from stochastic to nonlinear dynamics and beyond. *Journal of the Royal Society, Interface / the Royal Society*, 8(54):107–16, jan 2011.

[68] Melissa Vellela and Hong Qian. Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlögl model revisited. *Journal of the Royal Society, Interface / the Royal Society*, 6(39):925–40, oct 2009.

[69] J Keizer and R F Fox. Qualms regarding the range of validity of the glansdorff-prigogine criterion for stability of non-equilibrium States. *Proceedings of the National Academy of Sciences of the United States of America*, 71(1):192–6, jan 1974.

[70] P Glansdorff, G Nicolis, and I Prigogine. The thermodynamic stability theory of non-equilibrium States. *Proceedings of the National Academy of Sciences of the United States of America*, 71(1):197–9, jan 1974.

[71] Herbert B. Callen. Principle of Minimum Entropy Production. *Physical Review*, 105(2):360–365, jan 1957.

[72] Howard Salis and Yiannis Kaznessis. Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions. *The Journal of Chemical Physics*, 122(5):054103, feb 2005.

[73] Howard Salis and Yiannis N. Kaznessis. An equation-free probabilistic steady-state approximation: Dynamic application to the stochastic simulation of biochemical reaction networks. *The Journal of Chemical Physics*, 123(21):214106, dec 2005.

[74] Olegh Bilous and Neal R. Amundson. Chemical reactor stability and sensitivity. *AIChE Journal*, 1(4):513–521, dec 1955.

[75] Rutherford Aris and Neal R. Amundson. An analysis of chemical reactor stability and control—I. *Chemical Engineering Science*, 7(3):121–131, 1958.

[76] R.A. Schmitz and N.R. Amundson. An analysis of chemical reactor stability and control—Vb. *Chemical Engineering Science*, 18(7):391–414, jul 1963.

[77] K.S. Kumar and R. Sridhar. On the identification of control systems by the quasi-linearization method. *IEEE Transactions on Automatic Control*, 9(2):151–154, apr 1964.

[78] Ross Taylor. Solution of the linearized equations of multicomponent mass transfer. *Industrial & Engineering Chemistry Fundamentals*, 21(4):407–413, nov 1982.

[79] V. I. Yudovich. *The linearization method in hydrodynamical stability theory*. Translations of Mathematical Monographs, American Mathematical Society, 1989.

[80] Tor Steinar Schei. A finite-difference method for linearization in nonlinear estimation algorithms. *Automatica*, 33(11):2053–2058, nov 1997.

[81] Jonathan W. Wittwer, Kenneth W. Chase, and Larry L. Howell. The direct linearization method applied to position error in kinematic linkages. *Mechanism and Machine Theory*, 39(7):681–693, jul 2004.

[82] Olegh Bilous and Neal R. Amundson. Chemical reactor stability and sensitivity: II. Effect of parameters on sensitivity of empty tubular reactors. *AIChE Journal*, 2(1):117–126, 1956.

[83] B Wayne Bequette. Nonlinear Control of Chemical Processes: A Review. *Znd. Eng. Chem. Res*, 30:1391–1413, 1991.

[84] Geoffrey M. Cooper and Robert E. Hausman. *The Cell, A molecular Approach*. Sinauer Associates, Inc., Sunderland, Massachusetts, sixth edition, 2013.

[85] Brian Munsky and Mustafa Khammash. The finite state projection algorithm for the solution of the chemical master equation. *The Journal of Chemical Physics*, 124(4):044104, jan 2006.

[86] Tobias Sutter, David Sutter, Peyman Mohajerin Esfahani, and John Lygeros. Generalized maximum entropy estimation. 2017.

[87] J.R. Dormand and P.J. Prince. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6(1):19–26, mar 1980.

[88] R. Lefever and G. Nicolis. Chemical instabilities and sustained oscillations. *Journal of Theoretical Biology*, 30(2):267–284, feb 1971.

[89] B. Lavenda, G. Nicolis, and M. Herschkowitz-Kaufman. Chemical instabilities and relaxation oscillations. *Journal of Theoretical Biology*, 32(2):283–292, aug 1971.

[90] John Goutsias. Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. *The Journal of Chemical Physics*, 122(18):184102, may 2005.

[91] J. C. Butcher. A stability property of implicit Runge-Kutta methods. *BIT*, 15(4):358–361, dec 1975.

[92] Joakim Munkhammar, Lars Mattsson, and Jesper Rydén. Polynomial probability distribution estimation using the method of moments. *PLOS ONE*, 12(4):e0174573, apr 2017.

[93] Pedro H. Constantino and Yiannis N. Kaznessis. Maximum entropy prediction of non-equilibrium stationary distributions for stochastic reaction networks with oscillatory dynamics. *Chemical Engineering Science*, 171:139–148, 2017.

[94] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, mar 1951.

[95] Slaven Peleš, Brian Munsky, and Mustafa Khammash. Reduction and solution of the chemical master equation using time scale separation and finite state projection. *The Journal of Chemical Physics*, 125(20):204104, nov 2006.

[96] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, fourth edition, 2013.

Appendix A

A.1 ZI Closure Scheme Accuracy

The most popular method to solve stochastic reaction networks is Gillespie's Stochastic Simulation Algorithm (SSA) [9]. Even though SSA is extremely accurate, it is computationally expensive. As a result, SSA is often impractical for parametric analysis. For example, for the Schlögl model for the kinetic parameters reported in Chapter 4 for volume 3, SSA needs more than 1500 CPU hours to reach steady state for a simulation with 100,000 trajectories. In comparison, ZI closure scheme for the same system needs only 2 CPU seconds. ZI closure scheme is so much faster because it directly calculates the steady state distribution without calculating the transient behavior of the system. Additionally, ZI closure scheme computational time scales linearly with volume (figure A.1), which ensures that the computational time needed will stay in manageable sizes for a wide range of volumes.

Another way to calculate steady state distributions for the Schlögl reaction network (Chapter 4) is through the analytical solution [68]:

$$\frac{P(X)}{P(0)} = \prod_{n=0}^{X-1} \frac{\frac{k_2}{V}(n)(n-1) + k_4 V}{\frac{k_1}{V^2}(n+1)n(n-1) + k_3(n+1)} \quad , \quad P(0) = 1 - \sum_{k=0}^{\infty} P(k) \tag{A.1}$$

with $P(X)$ representing the stationary probability at the state space point X and k_i, $i = 1, \ldots, 4$, are the kinetic constants of the reaction network. $P(0)$ is the probability at state space point $X = 0$. The above equation can be easily derived from CME (equation (4.2)) if the probability flux in the forward direction is set equal to the probability flux in the backwards direction at each X at steady state.

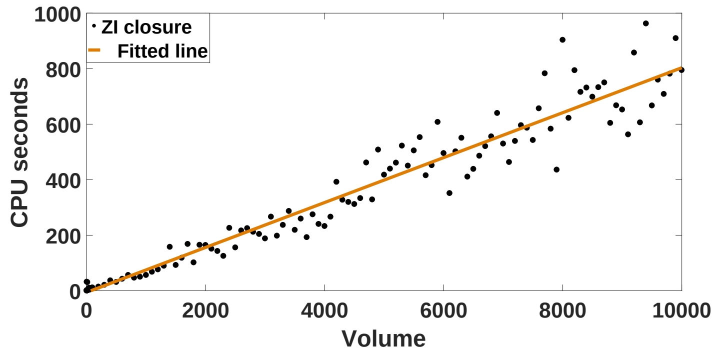


Figure A.1: CPU time required to run ZI closure scheme for different volumes. The results correspond for the Schlögl model presented in Chapter 4 for $k_4 = 4\ 10^4$.

The analytical solution appears to be as computationally efficient as ZI closure scheme. However, the method fails to produce results for large volume values (for our analysis this value is around volume $V = 40$). The numerical difficulty of the analytical solution lies on the need of calculating all the points divided by $P(0)$. The probability $P(0)$ goes to zero very fast as the system size increases. As a result, the ratio of probabilities then quickly goes to positive infinity (table A.1).

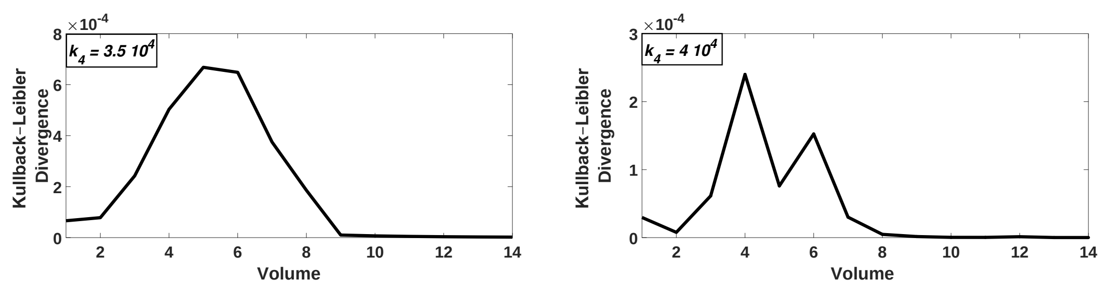
Table A.1: Second row presents representative values of the stationary probability distribution at $X = 0$ ($P(0)$) at different volumes (first row). The table also includes the ratio between the maximum value of the stationary probability distribution over $P(0)$. The values are obtained for the Schlögl of Chapter 4 for $k_4 = 4.0\ 10^4$.

Volume	1	5	10	15
$P(0)$	$1.89\ 10^{-7}$	$4.06\ 10^{-35}$	$4.14\ 10^{-70}$	$4.16\ 10^{-105}$
$\frac{\text{maximum } P(X)}{P(0)}$	$1.14\ 10^{5}$	$2.19\ 10^{32}$	$1.54\ 10^{67}$	$1.26\ 10^{102}$

In Chapter 4, we investigate reduced volume sizes that reach in the thousands. The largest number that Matlab can handle is $1.7977\ 10^{308}$, so we quickly get infinities in the results. To the author's knowledge, there is no way to compute the solution using the analytical expression for reduced volumes larger than 40, given the kinetic constant values the author employee.

Here also lies the significance of ZI closure scheme approach. It uses a probabilistic modeling formalism to study systems with sizes that span a vast range, practically tending to the thermodynamic limit and can thus investigate the transition between stochastic and deterministic dynamics in chemical reaction systems.

Importantly, ZI closure scheme is as accurate as the analytical solution, as observed in figure A.2. The accuracy of the method is not affected as the system size increases (figure A.2). This result creates confidence that ZI closure scheme produces accurate results even in areas that analytical solution is unable to produce results.



(a) (b)

Figure A.2: Difference between probability distributions calculated with analytical solution and ZI closure. The plot represents solutions for a variety of volumes for the Schlögl model. For the comparison, the Kullback - Leibler divergence was used [94]. Figure A.2a presents results for $k_4 = 3.5\ 10^4$ and figure A.2b for $k_4 = 4\ 10^4$.

A.2 Inverse of Jacobian

The Jacobian matrix J of the moment equations for a stochastic reaction network is calculated as:

$$J_{i,j} = \frac{\partial \mu_i}{\partial \lambda_j} = -\mu_{i+j} + \mu_i \mu_j \tag{A.2}$$

By the definition of probability moments (equation 6.4), it is true that:

$$\mu_{i+j} = \sum_{\Omega} \left\{ f_{\mu_{i+j}}(\boldsymbol{X}) \exp \left[-\sum_{j=0}^{\Psi} \lambda_j f_{\mu_j}(\boldsymbol{X}) \right] \right\} \tag{A.3}$$

Thus, $\mu_{i+j} = \mu_{j+i}$. It is then apparent that the Jacobian matrix is symmetric since:

$$J_{i,j} = -\mu_{i+j} + \mu_i \mu_j = -\mu_{j+i} + \mu_j \mu_i = J_{j,i} \tag{A.4}$$

As a symmetric matrix, J can be calculated as [96]:

$$J = Q^T \Lambda Q \tag{A.5}$$

where Λ is a diagonal matrix with the diagonal elements being the eigenvalues of J and Q the associated eigenvector matrix. Q is also real orthogonal [96], i.e. $Q^{-1} = Q^T$. The inverse of the Jacobian matrix, J^{-1} can be calculated as:

$$J^{-1} = \left(Q^T \Lambda Q\right)^{-1} = (Q)^{-1} (\Lambda)^{-1} \left(Q^T\right)^{-1} = Q^T (\Lambda)^{-1} Q \tag{A.6}$$

Since Λ is diagonal, $(\Lambda)^{-1}$ is also diagonal with elements the reciprocal elements of Λ ($(\Lambda_{i,j})^{-1} = \frac{1}{\Lambda_{i,j}}$). Additionally, each element of Λ, $\Lambda_{i,j}$, represents each eigenvalue of the Jacobian J.

Based on equation A.6, in order to calculate the inverse of the Jacobian matrix, one only needs to solve the eigenproblem of the Jacobian.

A.3 Runge-Kutta Tableau

In this section we list the Runge-Kutta RK5(4)7M tableau that is used in Chapter 6. The method was first described by Dormand and Prince [87] and we just list it here for the convenience of the reader. The Runge-Kutta tableau of the method is:

		$j = 1$	2	3	4	5	6		
i	c_i	a_{ij}						$\hat{b}_i$	b_i
1	0							$\frac{35}{384}$	$\frac{5179}{57600}$
2	$\frac{1}{5}$	$\frac{1}{5}$						0	0
3	$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$					$\frac{500}{1113}$	$\frac{7571}{16695}$
4	$\frac{4}{5}$	$\frac{44}{45}$	$-\frac{56}{15}$	$\frac{32}{9}$				$\frac{125}{192}$	$\frac{393}{640}$
5	$\frac{8}{9}$	$\frac{19372}{6561}$	$-\frac{25360}{2187}$	$\frac{64448}{6561}$	$-\frac{212}{729}$			$-\frac{2187}{6784}$	$-\frac{92097}{339200}$
6	1	$\frac{9017}{3168}$	$-\frac{355}{33}$	$\frac{46732}{5247}$	$\frac{49}{176}$	$-\frac{5103}{18656}$		$\frac{11}{84}$	$\frac{187}{2100}$
7	1	$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$	0	$\frac{1}{40}$