Incorporating Stochastic Influences in Assembly Models: Application to Intermediate Filament Polymerisation

Morgan Craig

A Thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of Master of Science

Department of Mathematics
University of Manitoba
Winnipeg

Copyright ©2011 by Morgan Craig
Abstract

The focus of this thesis is the inclusion of stochasticity into mathematical models of assembly with particular interest to the in vitro polymerisation of intermediate filaments, one of three components of the cytoskeleton. From the chemical master equation (CME), two additional models (the reaction rate equations or RREs and the two-moment approximation equations or 2MA equations) are derived. As analysis of the CME is generally intractable, we present the stochastic simulation algorithm (SSA) as a means of reproducing the most probable state of the CME at a given time. The results from the SSA are compared to simulations of both the RREs and the 2MA equations and we find that the three models are in good agreement. Further, the numerical results are compared to mean lengths and length distributions of experimental data which all models are shown to mimic. Mathematical analyses of the RREs demonstrate the conservation of mass in the system, and the unique positive equilibrium is proven to be globally asymptotically stable. Further, the 2MA equations are also shown to have conservation of mass and to possess an analogous equilibrium to the one found in the case of the RREs. In general, this study illustrates how randomness can be incorporated in polymerisation models and highlights the advantages and disadvantages of the different approaches.
Acknowledgements

I would like to first acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) for their support in the first year of my Master’s degree through the Postgraduate Scholarship (PGS) and the Michael Smith Foreign Study Supplement.

I wish to wholeheartedly express my gratitude to Prof. Dr. Michael Beil (Department of Internal Medicine, University of Ulm, Germany), Prof. Dr. sc. nat. ETH Othmar Marti (Experimental Physics, University of Ulm, Germany), and Prof. Dr. Harald Herrmann (Division of Molecular Genetics, German Cancer Research Center DKFZ, Heidelberg, Germany) for their gracious reception and support during my time in Germany. Their knowledge and expertise were invaluable to the present manuscript.

I also recognise the work of my examining committee, of which I am very appreciative. Last, I wish to sincerely thank my supervisor Dr. Stéphanie Portet for her indefatigable support and patience throughout my studies at the University of Manitoba, particularly in the preparation of this thesis during my final year. Her encouragement allowed me to grow as a student and researcher and for this I am forever indebted.
Dedication

To my parents, who raised me to question and examine. To my sister for all of her grammar expertise and willingness to read through my language foibles and to Isaac, Noah, and Xavier. To my friends, particularly Bryan, Haley, Jenn, Jesse, Rachel, Roberta, Shoni, and Stacey. You all kept me on task, grounded and sane, which was no small feat. Thanks. Finally to my Gramma Margaret and my Baba Ettie.
## Contents

1 Introduction
   1.1 The Cytoskeleton and Intermediate Filaments .................................. 1
   1.2 Motivation ......................................................................................... 6

2 Models
   2.1 Fundamentals .................................................................................. 14
      2.1.1 Total Number of Reactions With Filament of Length $i$ as Smallest Reactant .................................................. 16
      2.1.2 Total Number of System Reactions ............................................. 19
   2.2 Chemical Master Equation ............................................................... 21
      2.2.1 An Example ................................................................................ 26
      2.2.2 Simulating the CME: Stochastic Simulation Algorithm ............... 30
   2.3 Reaction Rate Equations .................................................................. 35
   2.4 Two Moment Approximation ............................................................ 41

3 Results
   3.1 Chemical Master Equation ............................................................... 55
      3.1.1 Necessary Number of Simulations .............................................. 63
      3.1.2 Effect of Initial Condition on SSA Output .................................. 70
      3.1.3 Effect of Parameter Changes on SSA Output ............................. 73
   3.2 Reaction Rate Equations .................................................................. 79
      3.2.1 Existence, Uniqueness and Positivity of Solutions ..................... 79
      3.2.2 Conservation of Mass .................................................................. 81
      3.2.3 Existence of Unique Equilibrium Point ..................................... 85
      3.2.4 Global Stability Analysis ............................................................ 86
   3.3 Two Moment Approximation ............................................................ 89
      3.3.1 Existence and Uniqueness of Solutions ..................................... 90
      3.3.2 Equilibrium Results ................................................................. 91
      3.3.3 Conservation of Mass and Notes on the 2MA Equations .......... 93
CONTENTS

3.4 Comparisons .............................................. 102
  3.4.1 Experimental Results .................................. 103
  3.4.2 Modelling Approaches .................................. 116
  3.4.3 Frameworks ............................................. 119

4 Discussion and Remarks 126
  4.1 Treatment of Mathematical Results ....................... 128
  4.2 Computational Approaches ................................. 130
  4.3 Experimental Comparison .................................. 133

5 Conclusion 136

Relevant Definitions and Theorems 139
  5.1 Vector Calculus and Matrix Theory ....................... 139
  5.2 Combinatorics and Number Theory ....................... 143
  5.3 Probability Theory ....................................... 144
  5.4 Theory of Dynamical Systems ............................ 149

Bibliography 152
List of Tables

1 Complete List of Operators, Parameters, and Variables . . . . ix

3.1 General SSA Results .................................................. 60
3.2 Effect of the Number of Simulations ................................. 66
3.3 Effect of the Number of Simulations ................................. 67
3.4 Effect of the Number of Simulations ................................. 69
3.5 Effect of Initial Condition ............................................. 71
3.6 Comparison of SSA and Experimental Data ....................... 106
3.7 Comparisons of Models to Experimental Data .................... 109
3.8 Model Comparison ...................................................... 117
List of Figures

1.1 Model Advantages .................................................. 10

2.1 Model Definitions .................................................. 13
2.2 Intermediate Filament Assembly Mechanisms ..................... 16

3.1 Analytic and Numerical Tools ...................................... 54
3.2 General SSA Results ............................................... 56
3.3 General SSA Results ............................................... 57
3.4 General SSA Results ............................................... 59
3.5 General SSA Results ............................................... 64
3.6 Effect of the Number of Simulations .............................. 68
3.7 Effect of Initial Condition ......................................... 72
3.8 Effect of Parameters ............................................... 75
3.9 Effect of Parameters ............................................... 76
3.10 Effect of Parameters ............................................... 77
3.11 Comparison of SSA and Experimental Data ....................... 107
3.12 Comparison of SSA and Experimental Data ....................... 110
3.13 Comparison of RREs and Experimental Data ...................... 111
3.14 Comparison of 2MA and Experimental Data ....................... 112
3.15 Comparison of Models to Experimental Data ...................... 115
3.16 Model Comparison .................................................. 118
3.17 Comparison of SSA to RRE ......................................... 121
3.18 Comparison of SSA to RRE ......................................... 122
3.19 Comparison of SSA to RRE ......................................... 123

4.1 Summary of Results .................................................. 127
<table>
<thead>
<tr>
<th>Variable</th>
<th>Domain</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_i$</td>
<td></td>
<td>Filament of length $i$ reacting in a chemical reaction</td>
</tr>
<tr>
<td>$N$</td>
<td>$N \in \mathbb{N}$</td>
<td>Total number of populations</td>
</tr>
<tr>
<td>$M$</td>
<td>$M \in \mathbb{N}$</td>
<td>Total number of reactions</td>
</tr>
<tr>
<td>$R_j$</td>
<td></td>
<td>Reaction $j$ ($j = 1, \ldots, M$)</td>
</tr>
<tr>
<td>$\mathbf{N}(t) = (N_1(t), N_2(t), \ldots, N_N(t))^t$</td>
<td>$\mathbf{N}(t) \in \mathbb{N}^N$</td>
<td>State vector: entries $N_i(t)$ are number of filaments of length $i$ ($i = 1, \ldots, N$)</td>
</tr>
<tr>
<td>$\mathbf{n}(t) = (n_1(t), n_2(t), \ldots, n_N(t))^t$</td>
<td>$\mathbf{n}(t) \in \mathbb{N}^N$</td>
<td>Sample vector: entries $n_i(t)$ are number of filaments of length $i$ as a sample of $\mathbf{N}(t)$ ($i = 1, \ldots, N$)</td>
</tr>
<tr>
<td>$\mathbf{X}(t) = (X_1(t), X_2(t), \ldots, X_N(t))^t$</td>
<td>$\mathbf{X}(t) \in \mathbb{R}^N$</td>
<td>State vector: entries $X_i(t)$ are concentrations of filaments of length $i$ ($i = 1, \ldots, N$)</td>
</tr>
<tr>
<td>$\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_N(t))^t$</td>
<td>$\mathbf{x}(t) \in \mathbb{R}^N$</td>
<td>Sample vector: entries $x_i(t)$ are number of filaments of length $i$ ($i = 1, \ldots, N$) as a sample of $\mathbf{X}(t)$</td>
</tr>
<tr>
<td>Variable</td>
<td>Domain</td>
<td>Explanation</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>( S )</td>
<td>( S \in \mathbb{Z}^{N \times M} )</td>
<td>Stochiometric matrix: entries ( S_{ij} ) describe the change in population ( i ) by occurrence of the reaction ( R_j )</td>
</tr>
<tr>
<td>( h(n) )</td>
<td>( h(n) \in \mathbb{N}^M )</td>
<td>Entries ( h_j(n) ) are the number of combinations of reactants of the reaction ( R_j ) ((j = 1, ..., M))</td>
</tr>
<tr>
<td>( C )</td>
<td>( C \in \mathbb{R}^M )</td>
<td>( C_j ) is the probability reaction rate of the reaction ( R_j ) ((j = 1, ..., M))</td>
</tr>
<tr>
<td>( a(n) )</td>
<td>( a(n) \in \mathbb{R}^M )</td>
<td>( a_j(x) = C_j h_j(x) ), ((j = 1, ..., M)) is the propensity function of the reaction ( R_j )</td>
</tr>
<tr>
<td>( P(n, t) )</td>
<td>( P(n, t) \in [0, 1] )</td>
<td>Probability of being in state ( n ) at time ( t )</td>
</tr>
<tr>
<td>( E[X_i] = \mu_i )</td>
<td>( E[X_i] \in \mathbb{R} )</td>
<td>Expectation of the random variable ( X_i )</td>
</tr>
<tr>
<td>( Cov(X_i, X_k) = \sigma_{ik} )</td>
<td>( Cov(X_i, X_k) \in \mathbb{R} )</td>
<td>Covariance of random variables ( X_i, X_k )</td>
</tr>
<tr>
<td>( A \odot B )</td>
<td>( A \odot B \in \mathbb{R}^{N \times M} )</td>
<td>Hadamard product of ( A, B ) ((A, B \in \mathbb{R}^{N \times M}))</td>
</tr>
<tr>
<td>( Df = \frac{\partial f(x)}{\partial x} )</td>
<td>( Df \in \mathbb{R}^{m \times n} )</td>
<td>Jacobian matrix of ( f(x) ), ( x \in \mathbb{R}^n, f \in \mathbb{R}^m )</td>
</tr>
<tr>
<td>( D^2 f = \frac{\partial^2 f(x)}{\partial x^2} )</td>
<td>( D^2 f \in \mathbb{R}^{n \times n} )</td>
<td>Hessian matrix of ( f(x) ), ( x \in \mathbb{R}^n, f : \mathbb{R}^n \to \mathbb{R} )</td>
</tr>
</tbody>
</table>

Table 1: Complete List of Operators, Parameters, and Variables
Chapter 1

Introduction

This thesis is concerned with mathematical models of assembly processes and the incorporation of stochastic effects on these systems with specific application to the in vitro polymerisation of intermediate filaments (IFs).

We begin by establishing the biological foundations of the IF system as the core of the present study. Subsequently, we detail the mathematical interest of this investigation and outline the format of this manuscript.

1.1 The Cytoskeleton and Intermediate Filaments

One of the principal motivations of the present research is the assembly of IFs, one of three components of the eukaryotic cell’s cytoskeleton. First hypothesised by Dujardin in 1835 [10, 77], the cell’s cytoskeleton is formed by three protein filaments: microfilaments (MFs) or actin, intermediate filaments (IFs), and microtubules (MTs), classified according to their diame-
tres (7nm, 10nm, and 25nm respectively). These filaments cross the internal structure of the cell and are linked in complex networks which are constantly assembling and disassembling to perform their cellular functions [89]. In eukaryotic cells the external layers (membranes or cell walls) are connected by the cytoskeleton to the nucleus, the central regulator of the cell comprising the genetic material of the organism. This framework of proteins not only gives structure to the cell, but also acts as a scaffold in the process of signal transduction [8,35]. Further, the cytoskeleton is also intimately involved during cellular division and is tasked with maintaining the organelles and certain cytoplasmic enzymes in place, contracting muscles cells, and providing motility to the cell [77]. Since Lowey first identified the presence of both actin and myosin in the slime mold in 1952 [53], the cytoskeleton has been studied in depth. In the following material, we will give a brief overview of MFs and MTs, and then provide a more comprehensive discussion of IFs.

MFs are formed of actin, the most abundant of all cellular proteins in nearly all eukaryotic cells [77]. Actin monomers are globular proteins. In the simplest model, actin filaments are formed of two wound chains of these globular proteins, imparting an overall polarity to the MF chain. In this context, polarity refers to an asymmetrical configuration at filament ends (plus and minus ends) resulting in a difference in electric charge at each terminus. As a result of the charged ends, the polymerisation of actin occurs at a fast-growing plus end and a slow-growing minus end [37]. The elongation of the actin filaments continues until an equilibrium equivalent to the critical concentration of monomers is reached. One can view the growth of MFs to be
occurring at the plus end by the single addition of actin monomers, while these subunits are removed one at a time at the minus end, in an event known as treadmilling. In fact, the rates of this association and dissociation have been determined in the case of MFs [70]. Actin filaments, together with myosin, are most notably crucial to muscular contractions, according to the sliding filament model of muscle contraction [39, 40], but MFs are also implicated in cellular motility and play the important role in the formation of the cleavage furrow during cytokinesis in animal cell mitosis [30].

Microtubules are also essential for cellular motility [77], as well as being a crucial regulator in the separation of the chromosome during mitosis [63]. According to [77], MTs are the most studied cytoskeletal component, owing to their mitotic involvement. They are hollow structures formed of tubulin protofilaments and, as is the case of actin filaments, MTs are polar structures [37]. Their assembly is regulated by microtubule-associated proteins [55] and centrosomes [63], the organisational center of the cell (in vivo assembly only). MT elongation occurs by the one-by-one addition of tubulin dimers at the plus end of the filament during a period of growth. This extension is then contrasted with a rapid phase of disassembly known as catastrophe at the minus end in a process similar to the treadmilling phenomenon present in actin [9, 77].

In contrast to the elongation by single subunits observed in the cases of MFs and MTs, the lengthening of intermediate filaments is not limited to the addition of single subunits; an IF can also assemble with filaments of
any length. The IF subunits are called *unit length filaments* or ULFs [32] which are formed by the lateral aggregation of dimers during a phase of rapid assembly. Due to the short duration of this period, the smallest building block in IF assembly is considered to be the ULF [46]. In fact, IFs differ from MFs and MTs in several ways, including their polarity (IFs are apolar structures [9]) and their mechanical properties (IFs are significantly more flexible than their cytoskeletal counterparts and act to hold the cell’s shape together) [8, 31, 43, 45, 47]. Intermediate filaments are classified into five subclasses, called *sequence homology classes*, or SHCs, according to their protein constituents [31, 68]:

- **SHC I**: acidic keratins
- **SHC II**: basic keratins
- **SHC III**: desmin/vimentin type proteins
- **SHC IV**: neurofilament proteins
- **SHC V**: nuclear lamina proteins
- **SHC VI**: lens proteins.

IF proteins are expressed in a cell depending on its level of differentiation, that is, for example, that IFs expressed in epithelial cells fall under SHCs I and II while those present in neurons are grouped in SHC IV.

IFs are also formed by a large class of genes. In fact, the genes encoding IFs are among the 100 largest gene families in the human genome [33, 68]. As
a result of their diversity and genetic variability, mutations on IFs are relatively more common and are not necessarily lethal to the cell. In contrast, mutations on MFs and MTs almost certainly result in cell death. Conditions caused by mutations on IFs proteins are thus able to present themselves in humans and are significantly heterogeneous. These diseases range from skin disorders arising from mutations on SHC I and II IFs, to Alexander’s disease, cardiomyopathy, and amyotrophic lateral sclerosis (SHC III proteins), to muscular dystrophy and progeria (SHC V nuclear IFs), and cataracts (lens proteins of class SHC VI). IF proteins are involved in no less than 80 human diseases [68]. Further, point mutations can alter the expression of certain IFs from one SHC to another, which is notably useful in the diagnosis of diseases. Generally speaking, mutations of IF proteins can interfere with their assembly, and thus affect their structure and function.

IF-precipitated diseases highlight the necessity of identifying and clarifying the precursors and principles of IF assembly and their functions [15,78]. Historically, the study of assembly was complicated by the in vitro insolubility of IFs: IFs very rarely disassemble in vitro [3]. However, the authors of [73] have successfully characterised the steps of the in vitro assembly of individual filaments for vimentin (SHC III). There nonetheless remains much work to be completed in the investigation of IF network assembly, beginning with the development of network measurement techniques [20] to the mathematical description of formation [3].

The biological consideration of this thesis is the in vitro assembly of in-
dividual vimentin IFs. While research to date has taken a completely deterministic approach [73], the present study is concerned with randomness as it pertains to the IF polymerisation process in order to answer the following questions:

1. **What is the influence of stochasticity on IF assembly?**

2. **How do we include these stochastic effects in mathematical descriptions of the elongation of IFs?**

### 1.2 Motivation

Mathematical modelling of cellular systems is a powerful investigatory tool used to understand processes and mechanisms occurring within the cell. In conjunction with laboratory experiments, mathematical models can be used to validate hypotheses, estimate experimental parameters, validate sets of data, and/or provide entirely new insights into the inner workings of the cell [2, 65]. Modelling in cellular biology necessitates an inherently interdisciplinary approach and allows for a more exhaustive study of the cell, its organelles, and their behaviours [87]. Previous studies of cytoskeletal assembly focused primarily on the IFs counterparts: microfilaments (MFs) and microtubules (MTs). These investigations led to fairly complete pictures of the assembly mechanisms of MFs and MTs [35, 55, 63]. In addition, experimental results are reflected in mathematical models, for example, of the branching networks typical of MFs [11] and of the assembly kinetics of both MFs and MTs in [83].
Despite the noted relative lack of research on IFs, recent experimental progress has provided a great deal of insight into how IFs assemble. In addition, the in vitro experimental results are supported by an assembly model [73]. This model uses Smoluchowski’s coagulation equation [29] and a completely deterministic formalism to study the time evolution of the system of in vitro IFs, which in turn provides insight into the early behaviour of IF polymerisation. The mass-action model of [73] has been shown to be of a good fit to the data. However, its uniquely deterministic nature but approximates the mean of the system and, therefore, fails to account for any inherent randomness. Noise in biological systems can have little influence on the mean behaviours (for example, under conditions including large population levels and frequently occurring reactions) but it may also greatly affect the average comportment (see, for example, [34, 82, 86]). Further, the nonlinear nature of cellular systems can significantly alter the behaviours of a stochastic model from those of the deterministic analogue [14]. To have a complete understanding of IF assembly, we must study the effects of random behaviours within the system.

To begin to comprehend the noise present in the in vitro assembly of IFs, we must first understand from where it is derived. Loosely defined, statistical noise is the unexpected variation of the system. It is random in nature and therefore the study of a system’s noise must account for this randomness. Accordingly, the focus of this thesis is the derivation and study of two models of stochasticity known as the chemical master equation (CME) and the two-moment approximation (2MA) equations. Specifically, we examined their
application to the IF system in an attempt to evaluate the transient and asymptotic behaviour of IF assembly. We are further interested in the comparison of these models to their wholly deterministic counterparts, known as the reaction rate equations (RREs), which are shown to be analogous to the equations of [73]. Ultimately, the goal is to provide a comprehensive picture of IF assembly through a combination of modelling and experimental work.

The work of this thesis is divided as follows:

- Chapter 2 situates the in vitro assembly system of IFs. We then introduce the CME (Section 2.2), provide an illustrative example by means of the IF system, and outline an algorithm for its numerical simulation (Section 2.2.2). From the CME, we next construct the RREs (Section 2.3) with specific consideration given to reactions of zeroth and first order, and derive the 2MA equations (Section 2.4).

- Chapter 3 describes the numerical and mathematical results in the investigation of IF assembly, beginning again with the CME (Section 3.1), and following with the RREs (Section 3.2) and the 2MA equations (Section 3.3). Within these results, we also compare all three models to the experimental results in Section 3.4.1, as well as contrasting the different approaches (Sections 3.4.2 and 3.4.3).

- We conclude with discussions on the modelling approaches employed herein (Chapter 4) and offer suggestions for potential future work (Chapter 5).
Pertinent definitions and theorems are provided in the appendix ‘Relevant Definitions and Theorems’.

From the preceding material, it is becoming evident that the CME, the RREs, and the 2MA equations each offer advantages and disadvantages in the study of assembly models. As introductory evidence for these similarities and contrasts, consider Figure 1.1, which relates the correspondences between the three models. The present study will endeavour to highlight the benefits of each approach and to shed some light on their potential applications within mathematical cell biology.
Dynamics governing the time-evolution of probability distributions (purely stochastic)

Dynamics governing the means (fully deterministic)

Dynamics of means accounting for the system’s variability through covariances (deterministic with variability)

Figure 1.1: Connections of the chemical master equation (CME), the reaction rate equations (RREs), and the two-moment approximation (2MA) equations. Each arrow indicates how the passage from one model to another offers different information of the system in question. The indicated parallels and differences support the notion of a multi-pronged approach to the modelling of cellular systems.
Chapter 2

Models

This chapter discusses the fundamental concepts and derivations of the three different methods of study of intermediate filament (IF) assembly. Since we are interested in studying the length distributions of the IF system, we begin with a description of the system of IF polymerisation. Next we introduce the models used to translate this description mathematically. Throughout, we contrast the methods, and provide explanations of their utility while chronicling certain drawbacks inherent to each approach.

This chapter is structured as follows

1. Definition of the in vitro assembly of IFs and the general set-up of the biological system.

2. Introduction to the chemical master equation (CME).

3. Development of the reaction rate equations (RREs) from the CME.
4. Description of the two-moment approximation (2MA) equations by truncation of the CME’s probability distribution.

As will be detailed, the CME and the 2MA equations have the advantage of offering a study of the system’s noise, where the traditional RREs provide information on average behaviours without random effects. That being said, under certain assumptions, the use of the RREs is indeed validated. Figure 2.1 summarises the general definition of each of these approaches, and outlines the characteristics unique to each model.
Figure 2.1: Definitions and descriptions of the chemical master equation (CME), the two-moment approximation (2MA) equations, and the reaction rate equations (RREs). While each is a set of ODEs, they nonetheless offer different modelling strengths. Notably, the CME and the 2MA equations are able to capture random behaviours, while the RREs are not.
2.1 Fundamentals

In the present study, we consider a reaction system of fixed volume $\Omega$ with $N$ types of molecular populations. These molecules associate with each other via $M$ distinct reaction channels, and the occurrence of one such reaction induces changes to the copy number (or number of molecules) of the implicated populations.

Define

$$\mathbf{N}(t) = (N_1(t), N_2(t), ..., N_N(t))^T$$

to be the vector of dimension $N \times 1$ where each element $N_i(t)$ is a stochastic process accounting for the copy number of the molecule $i$ at time $t$. Drawing from probability theory, note that the expression

$$N_i(t) = n_i(t)$$

is interpreted to mean that the $i^{th}$ species has copy number $n_i$ at time $t$ and we define the state vector $\mathbf{n}(t)$ having elements $n_i(t)$ as the vector with dimensions $N \times 1$ which is a sample of the realisations of $\mathbf{N}(t)$.

If we are interested by molecular concentrations, we simply scale the copy numbers by the system’s volume. So $\mathbf{X}(t) = \frac{\mathbf{N}(t)}{\Omega}$ is the $N \times 1$-dimensional vector with elements
\[
X(t) = \left( X_1(t) = \frac{N_1(t)}{\Omega}, X_2(t) = \frac{N_2(t)}{\Omega}, \ldots, X_N(t) = \frac{N_N(t)}{\Omega} \right)^T
\]

and, similarly, the \(N \times 1\)-dimensional vector \(x(t)\) with elements \(x_i(t) = \frac{n_i(t)}{\Omega}\) is a sample of the stochastic process \(X(t)\). For simplicity, we will omit the dependence on \(t\), that is \(X(t) = X\), and similarly for all other variables. Table 1 summarises the above definitions. Further, as a convention, we define an \(N\)-vector to be a row vector of length \(N\). An \(N\)-element column vector will be referred to as an \(N \times 1\)-dimensional vector.

Before undertaking the derivation of the three models of interest in this thesis, we situate the biological problem that is foundational to the present study, namely the elongation of intermediate filaments by the longitudinal annealing from unit length filaments (ULF) as subunits. Accordingly, the length of a filament is measured in the number of ULFs. As is consistent with experimental results of [73], three mechanisms drive IF assembly:

1. ULF with ULF, giving a filament of length 2,
2. ULF with filament of length \(i\), giving a filament of length \(i + 1\),
3. Filament of length \(i\) with filament of length \(k\), giving a filament of length \(i + k\).

Note that a ULF is considered a filament of length 1. Interactions 2 and 3 are illustrated in Figure 2.2.
Throughout the modelling process, it is important to identify the number of reactions involving a filament of given length and the total number of reactions in the IF system, as in the following propositions.

### 2.1.1 Total Number of Reactions With Filament of Length $i$ as Smallest Reactant

For our purposes, we define two types of reaction mechanisms compatible with the three processes of IF assembly: *double* reactions, involving two filaments of the same length, and *mixed* reactions, involving two filaments of different lengths. Consider $F_i$ as a filament of length $i$ ($i = 1, \ldots, N$), interacting in a reaction system (the number of $F_i$ molecules at time $t$ is given by $N_i$). Then the double ($R_D$) and mixed ($R_M$) reactions can be viewed as the following chemical reactions

$$R_D : \quad 2F_i \rightarrow F_{2i}, \quad (2.1)$$

$$R_M : \quad F_i + F_k \rightarrow F_{i+k}. \quad (2.2)$$
These are the only two possible reactions in the systems because of the following hypotheses:

1. Assembly events occur between only two filaments at a time.

2. There is no disassembly in the system.

3. The experimental solution is well-mixed.

4. The reaction $F_i + F_k \rightarrow F_{i+k}$ is considered to be the same as $F_k + F_i \rightarrow F_{k+i}$. Therefore, as a convention, we consider reactions of type $F_i + F_k \rightarrow F_{i+k}$, where $i \leq k$.

Note that, as is consistent with experimental work, the reaction system begins with the introduction of $N$ ULFs, which can then be interpreted the total number of molecular populations in the system.

**Proposition 2.1.** The total number of potential reactions $\beta_i$ involving filaments of length $i$ as the shortest reactant length is equal to:

$$\beta_i = N - 2i + 1 \quad (2.3)$$

for $1 \leq i \leq \lfloor \frac{N}{2} \rfloor$, and $\beta_i = 0$ for $i > \lceil \frac{N}{2} \rceil$.

**Proof.** Let $N$ be the number of ULFs introduced. Since we assume there is no disassembly in the system and as ULFs are considered to be the building blocks of IF assembly, this implies the longest possible filament will be of length $N$. Let $\beta_i$ be the number of reactions implicating filaments of length $i$ ($F_i$) as the smallest reactant. We seek to show that
CHAPTER 2. MODELS

\[
\beta_i = \begin{cases} 
N - 2i + 1, & 1 \leq i \leq \left\lfloor \frac{N}{2} \right\rfloor \\
0, & \left\lfloor \frac{N+1}{2} \right\rfloor \leq i
\end{cases}
\]

By assumption, only two filaments at a time can interact in an assembly event. Fixing \( i \), we have that the number of filaments with which a filament \( F_i \) can interact is:

\[
\binom{N - i}{1} = N - i,
\]

as a filament \( F_N \) is the longest possible filament.

By convention, the reaction \( F_i + F_k \rightarrow F_{i+k} \) is considered to be the same as \( F_k + F_i \rightarrow F_{k+i} \), so we consider reactions to be of the form \( F_i + F_k \rightarrow F_{i+k} \) with \( i \leq k \). Note that there are \( i - 1 \) reactant lengths less than \( i \). Thus we discount the \( i - 1 \) interactions with filaments of length less than \( i \) from the reactions of Equation (2.4) and there are a total of

\[
\beta_i = N - i - (i - 1) = N - 2i + 1
\]

reactions implicating filaments of length \( i \) as the smallest reactant.

The bounds on \( i \) are given by the specific types of reaction in the system. Since we can form, at most, a filament of length \( N \), the longest of the smallest reactant in any reaction \( R_j \) will be \( i = \left\lfloor \frac{N}{2} \right\rfloor \). If \( N \) is even, the last
reaction is the double reaction

\[ 2F_{\frac{N}{2}} \rightarrow F_N. \]

Otherwise, \( i = \lfloor \frac{N}{2} \rfloor = \frac{N-1}{2} \) and the last reaction in the system is the mixed reaction given by:

\[ F_{\frac{N}{2} - \frac{1}{2}} + F_{N - \frac{N}{2} + \frac{1}{2}} \rightarrow F_N. \]

Hence

\[ \beta_i = N - 2i + 1 \]

for \( 1 \leq i \leq \lfloor \frac{N}{2} \rfloor \) and \( \beta_i = 0 \) otherwise, as required.

\[ \square \]

### 2.1.2 Total Number of System Reactions

To count the total number of reactions in the system, consider the following proposition.

**Proposition 2.2.** The total number of reactions \( M \) in the system is given by

\[ M = \sum_{i=1}^{\lfloor \frac{N}{2} \rfloor} (N - 2i + 1) = \sum_{i=1}^{N} \left\lfloor \frac{i}{2} \right\rfloor. \]  

(2.6)

**Proof.** We seek to show that

\[ \sum_{i=1}^{\lfloor \frac{N}{2} \rfloor} (N - 2i + 1) = \sum_{i=1}^{N} \left\lfloor \frac{i}{2} \right\rfloor. \]  

(2.7)
From Proposition 2.1,
\[
\beta_i = \begin{cases} 
N - 2i + 1, & 1 \leq i \leq \left\lfloor \frac{N}{2} \right\rfloor \\
0, & \left\lfloor \frac{N+1}{2} \right\rfloor \leq i
\end{cases}
\]
where \(\beta_i\) is the number of reactions involving filaments of length \(i\) as the shortest reactant length. Then considering all possible reactions, we have that
\[
\sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} \beta_i = \sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} (N - 2i + 1)
\]
\[
= \sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} N - 2\sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} i + \sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} 1
\]
\[
= \left\lfloor \frac{N}{2} \right\rfloor N - \left\lfloor \frac{N}{2} \right\rfloor \left( \left\lfloor \frac{N}{2} \right\rfloor + 1 \right) + \left\lfloor \frac{N}{2} \right\rfloor
\]
\[
= \left\lfloor \frac{N}{2} \right\rfloor \left( N - \left\lfloor \frac{N}{2} \right\rfloor \right). \tag{2.8}
\]
From [25], we have the result that
\[
\sum_{i=0}^{N} \left\lfloor \frac{i}{2} \right\rfloor = N \left\lfloor \frac{N}{2} \right\rfloor - \left( \frac{N}{2} \right)^2. \tag{2.9}
\]
Recognising that \(\left\lfloor \frac{0}{2} \right\rfloor = 0\), we can adjust the bounds of Equation (2.9) and we have that:
\[
\sum_{i=1}^{N} \left\lfloor \frac{i}{2} \right\rfloor = N \left\lfloor \frac{N}{2} \right\rfloor - \left( \frac{N}{2} \right)^2.
\]
CHAPTER 2. MODELS

This gives that:

$$\sum_{i=1}^{N} \left\lfloor \frac{i}{2} \right\rfloor = N \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{N}{2} \right\rfloor^2$$

$$= \left\lfloor \frac{N}{2} \right\rfloor \left( N - \left\lfloor \frac{N}{2} \right\rfloor \right), \quad (2.10)$$

and we have the equality of Equations (2.8) and (2.10) and Equation (2.7) holds. We have therefore that

$$M = \sum_{i=1}^{\lfloor \frac{N}{2} \rfloor} (N - 2i + 1) = \sum_{i=1}^{N} \left\lfloor \frac{i}{2} \right\rfloor.$$

\[\square\]

2.2 Chemical Master Equation

Having established the biological foundations of the IF system, we now turn to the construction of models to study the assembly of IFs. The chemical master equation (CME) is a set of coupled ordinary differential equations (ODEs) expressing the temporal evolution of the probabilities of all possible configurations in its state space [24]. The probability of interest $P(N(t) = n(t), t|N_0 = n_0, t_0)$ is defined as the probability that the system is in state $N = n$ at time $t$ given the initial state $N_0 = n_0$ at initial time $t_0$. For clarity, the notation of time-dependency has been omitted, as mentioned in Section 2.1. Further, in what follows, we write $P(n, t)$ to denote $P(N = n, t|N_0 = N_0, t_0)$. 
CHAPTER 2. MODELS

Consider the vector \( \mathbf{N} = (N_1, N_2, ..., N_N)^T \) of copy numbers (number of molecules), with sample \( \mathbf{n} = (n_1, n_2, ..., n_N)^T \), and the M reaction channels \( R_1, R_2, ..., R_M \), per Proposition 2.2. In what follows, we make use of the ensuing definitions.

**Definition 2.1.** Let the microscopic probability rate vector \( \mathbf{C} \) be an \( M \)-vector whose entries \( C_j \) give the specific probability rate constant with which the reaction \( R_j \) will occur.

By the law of mass-action for chemical reactions, the rate of any elementary reaction (as expressed in its rate equation) is proportional to the product of its reactant concentrations. The order of each reactant is the power of its concentration in this equation. With this in mind, we can define \( \mathbf{h}(\mathbf{n}) \).

**Definition 2.2.** Let \( \mathbf{h}(\mathbf{n}) \) be an \( M \)-vector whose entries \( h_j(\mathbf{n}) \) are the possible combinations of reactant molecules implicated in an \( R_j \) reaction given the system is in state \( \mathbf{n} \). To that end, let \( h_j(\mathbf{n}) \) be a mapping:

\[
h_j(\mathbf{n}) : \mathbb{N}^N \to \mathbb{N}
\]

such that

\[
h_j(\mathbf{n}) = \left( \frac{n_1}{O(n_1^j)} \right) \left( \frac{n_2}{O(n_2^j)} \right) \cdots \left( \frac{n_N}{O(n_N^j)} \right),
\]

where \( O(n_i^j) \) is the order of the \( i^{th} \) molecule in the \( j^{th} \) reaction.

Definitions 2.1 and 2.2 are used together to define the following reaction probabilities.
Definition 2.3. Let \( a_j(n) = C_j h_j(n) \), which is referred to as the propensity function of each reaction \( R_j \) (\( j = 1, \ldots, M \)). Each \( a_j(n) \) expresses the probability reaction \( R_j \) will take place. In vector form, let \( a(n) \) be the \( M \)-vector given by \( a(n) = C \circ h(n) \), where \( \circ \) is the Hadamard product (Definition 5.5).

Note that by Definition 2.3, the probability of an occurrence of the reaction \( R_j \) (\( j = 1, \ldots, M \)) within the time interval \([t, t + \tau]\) is given by

\[
a_j(n) \tau. \tag{2.11}
\]

Additionally, from Definition 2.2 we observe that the order of the molecule \( n_i (O(n^j_i)) \) is zero when it is not a reactant in the \( j^{th} \) reaction. Since \( \binom{n_i}{0} = 1 \), the contribution of \( n^j_i \) does not change the value of \( h_j(n) \). For example, let \( a_D(n) \) be the propensity function of a double reaction as in (2.1) and \( a_M(n) \) the propensity function of the mixed reaction as in (2.2) (with associated probability rates \( C_D \) and \( C_M \)). These propensity functions can then, in general, be expressed as:

\[
a_D(n) = C_D h_D = C_D \left( \binom{n_1}{O(n^D_1)} \cdots \binom{n_i}{O(n^D_i)} \cdots \binom{n_N}{O(n^D_N)} \right) \tag{2.12}
= C_D \cdot 1 \cdots 1 \cdot \binom{n_i}{2} \cdot 1 \cdots 1 \quad \text{(since } O(n^D_l) = 0, \forall l \neq i \text{)}
= C_D \frac{n_i(n_i - 1)}{2},
\]
\[ a_M(n) = C_M h_M \]
\[ = C_M \left( \frac{n_1}{O(n_1^M)} \right) \cdots \left( \frac{n_i}{O(n_i^M)} \right) \cdots \left( \frac{n_k}{O(n_k^M)} \right) \cdots \left( \frac{n_M}{O(n_M^M)} \right) \]
\[ = C_M \cdot 1 \cdots 1 \cdot \left( \frac{n_1}{1} \right) \cdots 1 \cdots \left( \frac{n_k}{1} \right) \cdots 1 \cdots 1 \] (since \( O(n_r^M) = 0, \forall r \neq i, k \))
\[ = C_M n_i n_k. \]

Finally, we introduce the matrix accounting for the state changes induced by each reaction.

**Definition 2.4.** Let \( S \) be an \( N \times M \) integer-valued matrix whose entries \( S_{ij} \) describe how the copy number of molecule \( i \) (\( i=1,\ldots,N \)) will be changed by the occurrence of an \( R_j \) reaction (\( j = 1,\ldots,M \)). The notation \( S \cdot j \) represents the column vector of the stochiometric coefficients of the reaction \( R_j \) for all molecules 1 to \( N \). Then \( S^T \) can be expressed as:

\[
S^T = \begin{pmatrix}
-2 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
-1 & -1 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\
-1 & 0 & -1 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & -2 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & -1 & -1 & 0 & 1 & 0 & 0 & \cdots & 0 \\
0 & -1 & 0 & -1 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & -2 & 0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & -1 & -1 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & -1 & 0 & -1 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & -2 & 0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\end{pmatrix}
\]

\[
\beta_1 = N - 1 \\
\beta_2 = N - 3 \\
\beta_3 = N - 5 \\
\beta_4 = \ldots
\]
Each $\beta_i$ is the total number of reactions implicating filaments of length $i$ as the smallest reactant. See Proposition 2.1 for the general expression of $\beta_i = N - 2i + 1$.

In addition, by the definition of $S$, we observed that when a filament of length $i$ is a reactant of the reaction $R_j$, its order is given by $O(n_j^i) = |S_{ij}|$. Since $S_{ij}$ accounts for the change in copy number of molecule $i$ in reaction $R_j$, when molecule $i$ reacts in a double reaction, $S_{ij} = -2 \Rightarrow |S_{ij}| = 2$. According to the law of mass action, this is precisely the order of molecule $i$ in the reaction rate equation of reaction (2.12). Similarly, if molecule $i$ is a reactant in a mixed reaction $R_j$, $|S_{ij}| = 1$ which is, again, the order of molecule $i$ in reaction $R_j$ (2.13). This, however, holds for reactant populations only, and is not generally true for products of reaction $R_j$.

Then, equipped with the above definitions, we have the CME given by:

$$
\frac{dP(n,t)}{dt} = \sum_{j=1}^{M} a_j(n - S_{.j})P(n - S_{.j}, t) - a_j(n)P(n, t).
$$

By collision theory, reactions occur only through the impact of reactant molecules in the proper orientation, with an underlying hypothesis that this impact produces a reaction with a given probability. The first terms of the right hand side of Equation (2.14) account for the probability of an occurrence of reaction $R_j$ within the infinitesimal interval $[t, t + \tau)$ to bring the system to the current state $n$. This implies that within $[t, t + \tau)$, the
reactants of $R_j$ have collided and, with probability $a_j(n - S_j)\tau$, $R_j$ can take place. Conversely, the second terms in the right hand side imply no reaction need occur to bring the system to state $n$ at time $t + \tau$, i.e. the system at time $t$ is already in state $n$.

**2.2.1 An Example**

To illustrate the CME given in Equation 2.14, consider the following example specific to the IF system. Recall that for IF polymerisation, we consider assembly events to occur between exactly two filaments at a time (e.g. a filament of length 6 could not be produced by the simultaneous reaction of filaments of lengths 1, 2, and 3). This, in turn, implies that the propensity functions $a_j(n)$ take a particular form. Begin by letting $N = 5$ be the total number of ULFs in the system. Since $\left\lfloor \frac{5}{2} \right\rfloor = 2$, we have from Proposition 2.1 that $\beta_1 = 5 - 2 + 1 = 4$ is the total number of reactions involving ULFs as the smallest reactant and $\beta_2 = 5 - 4 + 1 = 2$ is the total number of reactions implicating filaments of length 2 as the smallest reactant. By Proposition 2.2, the total number of reactions in the system is $M = \sum_{i=1}^{5} \left\lfloor \frac{i}{2} \right\rfloor = 6$. These six reactions are:

\[ R_1 : F_1 + F_1 \rightarrow F_2, \]
\[ R_2 : F_1 + F_2 \rightarrow F_3, \]
\[ R_3 : F_1 + F_3 \rightarrow F_4, \]
\[ R_4 : F_1 + F_4 \rightarrow F_5, \]
\[ R_5 : F_2 + F_2 \rightarrow F_4, \]
$R_6: F_2 + F_3 \rightarrow F_5$.

As outlined in Section 2.1, the copy number sample vector $n$ is given by

$$n = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{pmatrix}.$$

From the reactions detailed above, the transposed stochiometric matrix is defined to be:

$$S^T = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & -1 & 1 \\ 0 & -2 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \end{pmatrix},$$

which we observe to follow the general pattern defined in Definition 2.4.

In order to calculate the $M$-vector of propensity functions $a(n)$, we define the microscopic probability rate vector $C$:

$$C = \begin{pmatrix} C_1, & C_2, & C_3, & C_4, & C_5, & C_6 \end{pmatrix}.$$
CHAPTER 2. MODELS

From Definition 2.2 and the observation following Definition 2.4, the vector $\mathbf{h}(\mathbf{n})$ is computed as the number of combinations of the reactant populations for each reaction $R_j$, here with $j = 1, ..., 6$:

$$
\mathbf{h}(\mathbf{n}) = \left( \binom{n_1}{|S_{11}|}, \binom{n_1}{|S_{12}|} \binom{n_2}{|S_{22}|}, \binom{n_1}{|S_{13}|} \binom{n_3}{|S_{33}|}, \binom{n_1}{|S_{14}|} \binom{n_4}{|S_{44}|}, \binom{n_2}{|S_{25}|} \binom{n_3}{|S_{35}|}, \binom{n_2}{|S_{26}|} \binom{n_3}{|S_{36}|} \right)
$$

$$
\mathbf{h}(\mathbf{n}) = \left( \binom{n_1}{|S_{12}|}, \binom{n_1}{|S_{13}|}, \binom{n_1}{|S_{14}|}, \binom{n_2}{|S_{25}|}, \binom{n_2}{|S_{26}|} \right)
$$

$$
\mathbf{h}(\mathbf{n}) = \left( \frac{n_1(n_1-1)}{2}, n_1n_2, n_1n_3, n_1n_4, \frac{n_2(n_2-1)}{2}, n_2n_3 \right).
$$

So the vector of propensity functions $\mathbf{a}(\mathbf{n}) = \mathbf{C} \circ \mathbf{h}(\mathbf{n})$ is given as:

$$
\mathbf{a}(\mathbf{n}) = \left( C_1 \frac{n_1(n_1-1)}{2}, C_2n_1n_2, C_3n_1n_3, C_4n_1n_4, C_5 \frac{n_2(n_2-1)}{2}, C_6n_2n_3 \right).
$$

Finally, considering the states $\mathbf{n} - S_{-j}$, $j = 1, ..., 6$ (the system is not in state $\mathbf{n}$ at time $t$) and $\mathbf{n}$ (the system is already in state $\mathbf{n}$ at time $t$), we can express the CME of the IF assembly system in matrix form to be:

$$
\frac{d\mathbf{P}(\mathbf{n}, t)}{dt} = \begin{pmatrix}
C_1 \frac{(n_1+2)(n_1+1)}{2} \\
C_2(n_1 + 1)(n_2 + 1) \\
C_3(n_1 + 1)(n_3 + 1) \\
C_4(n_1 + 1)(n_4 + 1) \\
C_5 \frac{(n_2+2)(n_2+1)}{2} \\
C_6(n_2 + 1)(n_3 + 1)
\end{pmatrix}^T \begin{pmatrix}
P(n_1 + 2, n_2 - 1, n_3, n_4, n_5) \\
P(n_1 + 1, n_2 + 1, n_3 - 1, n_4, n_5) \\
P(n_1 + 1, n_2, n_3 + 1, n_4 - 1, n_5) \\
P(n_1 + 1, n_2, n_3, n_4 + 1, n_5 - 1) \\
P(n_1, n_2 + 2, n_3, n_4 - 1, n_5) \\
P(n_1, n_2 + 1, n_3 + 1, n_4, n_5 - 1)
\end{pmatrix}
$$
where the first terms of the right hand side express the probability of bringing the system into state $n$, and the second terms account for the probability the system is in state $n$, and will remain as such.

Generally speaking, if we were able to solve the CME, we would have a complete picture of the probability distribution of all possible states at any given time $t$ [34]. However the CME is defined on an uncountable or even infinite state space [42] and is virtually unsolvable but for simple systems with very few molecular populations. As analytical results are seemingly unobtainable for the IF assembly system, we must rely on numerical techniques to study the system’s behaviour. To that end, simulation algorithms based on the Monte-Carlo sampling method (using the generation of random numbers sampled from the probability density function of the variables of interest to approximate the solution to a given problem) have been developed [4, 61, 62]. One such numerical method is discussed in the following section.
2.2.2 Simulating the CME: Stochastic Simulation Algorithm

As referenced in Section 1.2, the CME can be used in an effort to account for the inherent random nature present in a given reaction system. In practice, however, the dimensions of the CME quickly make it intractable. The most widely used numerical method for its simulation is the stochastic simulation algorithm (SSA), or Gillespie’s algorithm, proposed by D.T. Gillespie in 1977 [21]. The SSA is a powerful algorithm; it provides information on every reaction occurring in the system, information that is equivalent to the actual probability distribution of the CME [34], and is also straightforward to code. In what follows, we discuss the derivation of the SSA from its foundation in the CME.

Unlike simulation techniques for ODEs, which use time step interpolation (fixed or variable) to approximate the value of solutions at \( t + \Delta t \) (\( \Delta t > 0 \)), the SSA works by exactly calculating the time to the next reaction in the system, and then computing which reaction this is. By allowing this reaction to take place, we can subsequently update the population copy numbers and then restart the process and calculate the time to and occurrence of the next reaction.

To begin, recall the CME given in Equation (2.14) and the definition of the propensity functions \( a_j(n) \) as the probability of an occurrence of \( R_j \), given the system’s condition is \( n \). Let \( \tau > 0 \) be a random variable (RV) accounting for the time to the occurrence of the next reaction [14]. Define
$P(\tau|n, t)$ as

$$P(\tau|n, t) = \text{probability no reaction occurs within } [t, t + \tau].$$

Then, noting that occurrences of reactions during distinct time intervals (take, for example, $[t, t + \tau]$ and $[t + \tau, t + \tau + d\tau]$) are independent (Definition 5.11), by Definition 5.10 we have:

$$P(\tau + d\tau|n, t) = \text{Probability of no reaction in time } [t, t + \tau)$$

$$\times \text{Probability of no reaction within } [t + \tau, t + \tau + d\tau)$$

$$= P(\tau|n, t) \times P(d\tau|n, t + \tau)$$

$$= P(\tau|n, t) \times \left(1 - \sum_{j=1}^{M} a_j(n)d\tau\right),$$

by Equation (2.11). Here, $\sum_{j=1}^{M} a_j(n)d\tau$ is the probability all reactions will occur within $[t + \tau, t + \tau + d\tau)$. Rewritten, the probability that no reaction occurs within $[t, t + \tau + d\tau)$ can be expressed as:

$$\frac{P(\tau + d\tau|n, t) - P(\tau|n, t)}{d\tau} = -\sum_{j=1}^{M} a_j(n)P(\tau|n, t)$$

which, as $d\tau \to 0$, is

$$\frac{dP(\tau|n, t)}{d\tau} = -\sum_{j=1}^{M} a_j(n)P(\tau|n, t).$$
Solving the above ODE for the initial condition $P(0|\mathbf{n}, t) = 1$ (that is that the probability that no reaction will occur at time $\tau = 0$ is 1), we obtain

$$P(\tau|\mathbf{n}, t) = e^{-\sum_{j=1}^{M} a_j(\mathbf{n})\tau}. \quad (2.15)$$

Next, define $P(\tau, j|\mathbf{n}, t)d\tau$ to be the probability that reaction $R_j$ will occur within $[t + \tau, t + \tau + d\tau)$, with $j$ the RV of the next reaction’s index. In plain language, we can express this probability as the probability no reaction occurs anytime within $[t, t + \tau)$ and the probability reaction $R_j$ takes place precisely within $[t + \tau, t + \tau + d\tau)$. This can therefore be expressed as

$$P(\tau, j|\mathbf{n}, t)d\tau = P(\tau|\mathbf{n}, t) \times a_j(\mathbf{n})d\tau. \quad (2.16)$$

Inserting Equation (2.15) into (2.16), we obtain

$$P(\tau, j|\mathbf{n}, t) = e^{-\sum_{j=1}^{M} a_j(\mathbf{n})\tau} \left( a_j(\mathbf{n}) \right) \frac{a_j(\mathbf{n})}{\sum_{j=1}^{M} a_j(\mathbf{n})} \sum_{j=1}^{M} a_j(\mathbf{n})e^{-\sum_{j=1}^{M} a_j(\mathbf{n})\tau}. \quad (2.17)$$

Then from $P(\tau, j|\mathbf{n}, t)d\tau$ we obtain Equation (2.17) which can be expressed as the joint density function of two RVs. Thus $P(\tau, j|\mathbf{n}, t)$ can be written as the product of the two following probability density functions (Definition 5.12) of the two independent RV (Definition 5.11) $j$ and $\tau$ [34]:

$\textit{CHAPTER 2. MODELS}$
1. Next reaction index \((j)\)

\[
\frac{a_j(n)}{\sum_{j=1}^{M} a_j(n)}, \quad (2.18)
\]

2. Time to next reaction \((\tau)\)

\[
\sum_{j=1}^{M} a_j(n) e^{-\sum_{j=1}^{M} a_j(n) \tau}. \quad (2.19)
\]

This means that the next reaction index \(j\) is a uniformly distributed RV (Definition 5.14) and that the time \(\tau\) to the next reaction is an exponentially distributed RV with expectation \(E[\tau] = \frac{1}{\sum_{j=1}^{M} a_j(n)}\) (Definition 5.15). Using the formulation given above the SSA will be used to provide the time to the next reaction and to indicate the exact reaction to take place to calculate an exact solution to the CME. The SSA relies upon random sampling techniques like the inverse transform sampling method to generate pseudorandom numbers. The inverse transform sampling technique works by randomly generating a uniformly distributed number (Definition 5.14; \(u \sim U[0,1]\)) then, given a cumulative distribution function \(F\) (Definition 5.13), computes the value of a random variable \(x\) by setting \(x = F^{-1}(u)\) [23]. Then, with this generation of pseudorandom numbers in mind, the formulation of the SSA is summarised below.

1. Initialise the system \((t=0)\) and input the following parameters: \(C, S, \) and \(tstop\), where \(tstop\) is the experimental time of interest.

2. Sample two uniformly distributed random numbers \(r_1\) and \(r_2\) on the
CHAPTER 2. MODELS

unit interval \([0, 1]\) \((r_1, r_2 \sim U[0, 1])\).

3. Calculate \(\tau\), the time to the next reaction. As \(\tau\) is exponentially distributed (Equation (2.19)), using inverse transform sampling, we have that \(r_1 = 1 - e^{-a_0(n)}\), with \(a_0(n) = \sum_{j=1}^{M} a_j(n)\). Hence

\[
\tau = -\ln \left(1 - r_1\right) \cdot \frac{1}{a_0(n)}.
\]

Moreover, since \(r_1 \sim U[0, 1]\), then \(1 - r_1 \sim U[0, 1]\). Hence calculate \(\tau\) as \(\tau = \frac{1}{a_0(n)} \ln \left(\frac{1}{r_1}\right)\).

4. Calculate the next reaction to fire \((R_j)\) by comparing the sums:

\[
\sum_{l=1}^{j-1} a_l(n) < r_2 a_0(n) \leq a_0(n).
\]

This inequality is a consequence of the uniform distribution of \(j\) given in (2.18). Let \(r_2 = \frac{a_j(n)}{\sum_{j=1}^{M} a_j(n)} = \frac{a_j(n)}{a_0(n)}\). Since we are interested with calculating the first reaction to fire within \([t, t + \tau]\), the next reaction index is the first such \(j\) satisfying \(\sum_{l=1}^{j-1} a_l(n) < r_2 a_0(n)\), but without exceeding \(a_0(n)\) (the maximal value, as it is the sum of all propensity functions).

5. Let the reaction \(R_j\) fire and update the time counter with the value of \(\tau\), so \(t \leftarrow t + \tau\). Similarly, update the populations as \(N \leftarrow N + S_{.j}\) according to the \(R_j\) reaction that occurred. Return to Step 2.

6. Repeat Steps 2 through 6 until \(t > t_{stop}\).
As described above, the SSA allows us to reconstruct the possible states of the CME. However, despite the advantages offered by the CME and the SSA, the SSA approach can be time consuming and offers no analytical results \[13, 16\]. How then can one incorporate elements of stochasticity into a more analytically feasible system? Turning to the classical mass-action approach to reaction systems, we can study the temporal evolution of populations (instead of probabilities as is the case for the CME). Section 2.3 will show how a similar approach is well suited to linear reactions, but how nonlinear reactions necessitate the inclusion of the system’s noise. For this, we introduce and derive the two-moment approximation (2MA) equations \[84\] in Section 2.4.

### 2.3 Reaction Rate Equations

As noted, the classical approach to reaction systems is to employ a purely deterministic model based on the law of mass action kinetics which states, under the hypothesis of a well-stirred solution, that the rate of a reaction is proportional to the product of the reactant molecules concentrations (as expressed by the law of mass-action). These models account for macroscopic changes of the molecular concentration levels by studying their temporal evolution and are supported by the hypotheses of high molecular concentrations and frequently occurring reactions. Under these hypotheses, we are reasonably assured of obtaining an accurate portrayal of the average behaviour of the populations without considering any randomness and have more readily accessible tools for analysis (as shown in Figure 3.1(a)). In
this section we discuss the reaction rate equations (RREs) and their derivation from the CME.

Given the probability $P(n,t)$ from the CME of Equation (2.14), applying Theorem 5.2 with $T(N) = N_i$ (as a polynomial, $T(N)$ satisfies the conditions of Theorem 5.2), we have that

$$
\sum_n n_i \frac{dP(n,t)}{dt} = \sum_{j=1}^{M} E[((N_i + S_{ij}) - N_i)a_j(N)]
$$

$$
= \sum_{j=1}^{M} E[(N_i + S_{ij})a_j(N)] - E[N_i a_j(N)]
$$

$$
= \sum_{j=1}^{M} \sum_n [(n_i + S_{ij})a_j(n)P(n,t) - n_i a_j(n)P(n,t)]
$$

$$
= \sum_{j=1}^{M} \sum_n S_{ij}a_j(n)P(n,t)
$$

$$
= \sum_{j=1}^{M} E[S_{ij}a_j(N)]
$$

$$
= E \left[ \sum_{j=1}^{M} S_{ij}a_j(N) \right],
$$

(2.20)

by the linearity of the expectation operator (Definition 5.16).

Observe that zeroth order reactions of the type

$$
C_0 \rightarrow F_i
$$
have constant propensity functions given by \( a_j(n) = C_j \), implying Equation (2.20) is given simply by \( \sum_n n_i \frac{dP(n,t)}{dt} = \sum_{j=1}^{M} S_{ij} C_j \).

First order reactions are expressed as

\[
F_i \rightarrow F_k,
\]

and have linear propensity functions \( a_j(n) = C_j n_i \). Accordingly, considering linear reactions (reactions of order 1), Equation (2.20) becomes

\[
\sum_n n_i \frac{dP(n,t)}{dt} = E \left[ \sum_{j=1}^{M} S_{ij} a_j(N_i) \right] = \sum_{j=1}^{M} S_{ij} C_j E[N_i]. \tag{2.21}
\]

Scaling the LHS of Equation (2.21) by \( \Omega \) (as \( x = \frac{n}{\Omega} \)), we have by Definition 5.16 that

\[
\frac{1}{\Omega} \sum_n n_i \frac{dP(n,t)}{dt} = \frac{1}{\Omega} \frac{dE[N_i]}{dt} = \frac{dE \left[ \frac{N_i}{\Omega} \right]}{dt} = \frac{d\mu_i}{dt}.
\]

From now on, \( \mu_i \) is defined to be the mean concentration of molecule \( i \), \( \mu_i = E[X_i] \). Equation (2.21) then becomes
\[
\frac{d\mu_i}{dt} = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} C_j E[N_i]
\]

\[
= \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} C_j \Omega \mu_i
\]

\[
= \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega \mu)
\]

\[
= f_i(\mu),
\]

where we define the drift functions as:

\[
f_i(x) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega x), \quad (2.22)
\]

with \(x\) is in concentration.

In the present study,

\[
\frac{d\mu_i}{dt} = f_i(\mu), \quad i = 1, ..., N \quad (2.23)
\]

are called the reaction rate equations (RREs), where the \(f_i(\mu)\) defined in Equation 2.22 are linear combinations of the propensity functions of the system. As the RREs are completely deterministic for a given IC, the system’s mean behaviour is entirely predictable and is not influenced by randomness [59, 86]. As previously shown, for zeroth and first order reactions the time evolution of the means \(\mu_i\) is given exactly by Equation (2.23) since the propensity functions are linear and \(E[f_i(\mu)] = f_i(\mu)\) [84]. This assumption
of linearity, however, is not valid on the whole for large classes of reaction systems.

Consider now the second order reaction

$$2F_i \xrightleftharpoons{\kappa_{2i}} F_{2i}, \quad (2.24)$$

where $\kappa_{2i}$ is the macroscopic rate of reaction. The usual mass action ODE describing this double reaction is given by

$$\frac{d\mu_i}{dt} = -\kappa_{2i}\mu_i^2, \quad (2.25)$$

as described by the coagulation equation. In general, the Smoluchowski coagulation equation [29] is defined to be

$$\frac{d\mu_k}{dt} = \frac{1}{2} \sum_{j=1}^{k-1} K_{j,k-j}\mu_j\mu_{k-j} - \mu_k \sum_{j=1}^{\infty} K_{k,j}\mu_j, \quad \text{for } k = 1, 2, ...$$

where $\mu_k$ is the concentration of particles of size $k$ and $K(x,y)$ is the coagulation kernel (rate of coagulation of particles of size $x$ with particles of size $y$). In the case of the same double reaction, the propensity function $a_j(\Omega\mu)$ is given by

$$a_j(\Omega\mu) = C_j\Omega\mu_i(\Omega\mu_i - 1),$$

with $C_j$ to be the microscopic probability reaction rate of $R_j$. Then, in the case of the RREs given in Equation (2.23), the temporal evolution of $\mu_i$ follows
if there is only this double reaction taking place in the system. It is clear that the RRE expression for double reactions given in (2.26) differs from the mass-action description of the coagulation equation of Equation (2.25).

As detailed above, when reactions are higher order (degree 2 or more), their propensity functions are nonlinear. From Equations (2.20) and (2.22), we have that

$$\frac{d\mu_i}{dt} = \frac{1}{\Omega} E \left[ \sum_{j=1}^{M} S_{ij} a_j(\Omega X) \right] = E[f_i(\mu)].$$

In this case, the expectation $E[a_j(\Omega X)]$ is not as readily deduced. Indeed, in the case any reaction rate is nonlinear (second order or higher), the evolution of the means depends on higher order moments (Definition 5.17). Once reaction equations are nonlinear, the RREs of Equation (2.23) can only approximate this mean since $E[f_i(\mu)] \neq f_i(\mu)$. That being said, it is generally accepted that the RREs are a reasonable model of reaction systems under the hypotheses of high molecular concentrations and frequently occurring reactions [21, 59]. In the next section, we show how to estimate $E[a_j(\Omega X)]$ when the propensity functions $a_j(\cdot)$ are nonlinear.
CHAPTER 2. MODELS

2.4 Two Moment Approximation

To account for the random behaviours in a system, one could characterise its probability distribution via the CME. We will show how the use of a truncation of the system’s moments can be used as a means to incorporate stochasticity and hopefully render the mathematical analysis of the 2MA equations by means of results from ODE theory more accessible than in the case of the CME. Similar to the derivation of Equation (2.20), we will use the CME and Theorem 5.2 to obtain a set of ODEs regulating the mean behaviour of the system, while taking their intrinsic variability into account.

Recall Equation (2.20):

$$\frac{dE[N_i]}{dt} = E \left[ \sum_{j=1}^{M} S_{ij} a_j(N) \right].$$

As already noted, \( \frac{dE[N_i]}{dt} = \frac{d\mu_i}{dt} \), so dividing each side by \( \Omega \), we have

$$\frac{dE[N_i]}{dt} = E \left[ \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(N) \right].$$

\[\iff \frac{d\mu_i}{dt} = E \left[ \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega X) \right], \tag{2.27} \]

where \( \Omega X = N \). Then, by Equation (2.22), Equation (2.27) can be expressed as:
\[
\frac{d\mu_i}{dt} = E[f_i(X)].
\] 

(2.28)

By definition, the drift functions are sums of the propensity functions \(a_j(n)\). Keep in mind that the \(a_j(n)\) express the probability of each reaction \(R_j\) and depends on the number of combinations of reactant molecules. As chemical reactions are generally at most second order, the \(a_j(n)\) are polynomials of degree 2. Since the propensity functions are analytic according to Definition 5.4, and sums of analytic functions are also analytic (Theorem 5.1), we can use a Taylor series expansion about the mean \(\mu\) to find an expression of the expectation of the drift functions (Definition 5.3). Further, we can truncate these Taylor series to an approximation of second order.

Let \(\frac{\partial f_i}{\partial x^T}\) be the Jacobian of \(f_i\) (Definition 5.1) and \(\frac{\partial^2 f_i}{\partial x \partial x^T}\) be its Hessian matrix (Definition 5.2). An approximate expression, centred at \(\mu\), of the right hand term of Equation (2.28) is given by:

\[
E[f_i(X)] \approx f_i(\mu) + E\left[\frac{\partial f_i(\mu)}{\partial X^T}(X - \mu) + \frac{1}{2}(X - \mu)^T \frac{\partial^2 f_i(\mu)}{\partial X \partial X^T}(X - \mu)\right].
\]

Note that by Definition 2.22, the \(f_i(\mu) \in \mathbb{R}\). Hence, by the linearity of the expectation operator and by the fact that the expectation \(E[\frac{\partial f_i(\mu)}{\partial X^T}(X - \mu)] = 0\) as \(\mu\) is the expected value of \(X\), we have the following:

\[
E[f_i(X)] \approx f_i(\mu) + E\left[\frac{1}{2}(X - \mu)^T \frac{\partial^2 f_i(\mu)}{\partial X \partial X^T}(X - \mu)\right].
\]
By virtue of the fact that the drift functions are the sum of propensity functions of the reactions as in (2.12) and (2.13), \( f_i(x) \) are polynomials which are at most of degree 2. Therefore, the Hessians \( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \) are symmetric, constant \( N \times N \) matrices. Applying then the equivalent form of the expectation of a quadratic form (bilinear form) given in Definition 5.19, we obtain:

\[
E[f_i(X)] \approx f_i(\mu) + \frac{1}{2} tr \left( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \sigma \right),
\]

with \( tr(\cdot) \) the trace of a square matrix and \( \sigma \) the covariance matrix, with \( \sigma_{ik} = Cov(X_i, X_k) \). Since \( \sigma_{ik} = \sigma_{ki} \) by Definition 5.18, \( \sigma \) is symmetric. This recovers the ODEs describing the evolutions of the means:

\[
\frac{d\mu_i}{dt} = f_i(\mu) + \frac{1}{2} tr \left( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \sigma \right), \tag{2.29}
\]

with \( f_i(\mu) = \frac{1}{\pi} \sum_{j=1}^{M} S_{ij} a_j(\Omega \mu) \). The first term of the RHS of (2.29) describes the variation of the means, as is conventionally expressed by the RREs (Equation (2.23)). With the addition of the second term, it is clear that via the covariances, the equations governing the change in the means now take the stochastic variability of the system into account \[84\].

Since Equation (2.29) includes each covariance \( \sigma_{ik} \), an equation governing their temporal-evolution is needed. To derive the ODEs for the covariances, consider Theorem 5.2 and let \( T(N) = N_i N_k \) (which is a polynomial and thus satisfies the conditions of Theorem 5.2). We then obtain
\[ \sum_{n} n_i n_k \frac{dP(n, t)}{dt} = \sum_{j=1}^{M} E[(N_i + S_{ij})(N_k + S_{kj}) - N_i N_k] a_j(N) \]

\[ = \sum_{j=1}^{M} \sum_{n} ((n_i + S_{ij})(n_k + S_{kj}) - n_i n_k) a_j(n) P(n, t) \]

\[ = \sum_{j=1}^{M} \sum_{n} (n_i S_{kj} + n_k S_{ij} + S_{ij} S_{kj}) a_j(n) P(n, t) \]

\[ = \sum_{j=1}^{M} E[(N_i S_{kj} + N_k S_{ij} + S_{ij} S_{kj}) a_j(N)] \]

\[ = E \left[ \sum_{j=1}^{M} (N_i S_{kj} a_j(N) + N_k S_{ij} a_j(N) + S_{ij} S_{kj} a_j(N)) \right]. \] (2.30)

Since \( \sum_{n} n_i n_k \frac{dP(n, t)}{dt} = \frac{dE[N_i N_k]}{dt} \), we obtain

\[ \frac{dE[N_i N_k]}{dt} = E \left[ \sum_{j=1}^{M} (N_i S_{kj} a_j(N) + N_k S_{ij} a_j(N) + S_{ij} S_{kj} a_j(N)) \right]. \] (2.31)

Dividing through by \( \frac{1}{\Omega^2} \) gives

\[ \frac{dE \left[ \frac{N_i N_k}{\Omega^2} \right]}{dt} = E \left[ \frac{1}{\Omega^2} \sum_{j=1}^{M} (N_i S_{kj} a_j(N) + N_k S_{ij} a_j(N) + S_{ij} S_{kj} a_j(N)) \right]. \]

Recognising that

\[ E \left[ \frac{N_i N_k}{\Omega^2} \right] = E[X_i X_k], \]
CHAPTER 2. MODELS

let
\[ B_{ik}(x) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} S_{kj} a_j(\Omega x) \] (2.32)

(with \( x \) is in concentration) be the diffusion function, as is compatible with the fluctuation-dissipation theorem of statistical thermodynamics (see [1, 44, 84]). Equation (2.31) becomes

\[
\frac{dE[X_i X_k]}{dt} = E \left[ \frac{1}{\Omega} \sum_{j=1}^{M} (N_i S_{kj} a_j(N) + N_k S_{ij} a_j(N) + S_{ij} S_{kj} a_j(N)) \right]
\]
\[
= E \left[ X_i \frac{1}{\Omega} \sum_{j=1}^{M} S_{kj} a_j(\Omega X) \right] + E \left[ X_k \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega X) \right]
\]
\[
+ E \left[ \frac{1}{\Omega} \sum_{j=1}^{M} \frac{S_{ij} S_{kj} a_j(\Omega X)}{\Omega} \right]
\]
\[
= E[X_i f_k(X)] + E[X_k f_i(X)] + E \left[ \frac{1}{\Omega} B_{ik}(X) \right].
\] (2.33)

Note that by Definition 5.18, \( \sigma_{ik} = E[X_i X_k] - \mu_i \mu_k \), meaning the LHS of (2.33) becomes

\[
\frac{d(\sigma_{ik} + \mu_i \mu_k)}{dt} = \frac{d\sigma_{ik}}{dt} + \frac{d\mu_i}{dt} \mu_k + \mu_i \frac{d\mu_k}{dt}.
\]

Then Equation (2.33) can be expressed as

\[
\frac{d\sigma_{ik}}{dt} + \frac{d\mu_i}{dt} \mu_k + \mu_i \frac{d\mu_k}{dt} = E[X_i f_k(X)] + E[X_k f_i(X)] + E \left[ \frac{1}{\Omega} B_{ik}(X) \right].
\]
Rearranging, we obtain

\[ \frac{d\sigma_{ik}}{dt} = E[X_i f_k(X)] - \mu_i \frac{d\mu_k}{dt} + E[X_k f_i(X)] - \frac{d\mu_i}{dt} \mu_k + \frac{1}{\Omega} E[B_{ik}(X)], \]

which together with Equation (2.28) gives

\[ \frac{d\sigma_{ik}}{dt} = E[X_i f_k(X)] - \mu_i E[f_k(X)] + E[X_k f_i(X)] - E[f_i(X)] \mu_k + \frac{1}{\Omega} E[B_{ik}(X)]. \]

So, by linearity, we have:

\[ \frac{d\sigma_{ik}}{dt} = E[(X_i - \mu_i) f_k(X)] + E[(X_k - \mu_k) f_i(X)] + \frac{1}{\Omega} E[B_{ik}(X)]. \] (2.34)

Since each drift function is scalar valued, \( E[(X_i - \mu_i) f_k(X)] = E[f_k(X)](X_i - \mu_i) \) and \( E[(X_k - \mu_k) f_i(X)] = E[f_i(X)](X_k - \mu_k) \). Using a truncated Taylor series expansion about the mean \( \mu \) of the drift functions, we have the following approximate expressions:

\[ f_k(X)(X_i - \mu_i) \approx f_k(\mu)(X_i - \mu_i) + \frac{\partial f_k(\mu)}{\partial X} (X - \mu)(X_i - \mu_i), \] (2.35a)

\[ f_i(X)(X_k - \mu_k) \approx f_i(\mu)(X_k - \mu_k) + \frac{\partial f_i(\mu)}{\partial X} (X - \mu)(X_k - \mu_k). \] (2.35b)

Note that the expectation \( E[f_k(\mu)(X_i - \mu_i)] = 0 \) in Equation (2.35a) and similarly for like terms in Equations (2.35b). Taking the term-wise expectation of (2.35a) and (2.35b), leaves us with:
CHAPTER 2. MODELS

\[ E[f_k(X)(X_i - \mu_i)] \approx E\left[ \frac{\partial f_k(\mu)}{\partial X^T}(X - \mu)(X_i - \mu_i) \right], \]

\[ E[f_i(X)(X_k - \mu_k)] \approx E\left[ \frac{\partial f_i(\mu)}{\partial X^T}(X - \mu)(X_k - \mu_k) \right]. \]

Let \( \sigma_i \) be the \( i^{th} \) row and \( \sigma_k \) be the \( k^{th} \) column of \( \sigma \). By convention, we will always consider \( i^{th} \) rows and \( k^{th} \) columns. As \( \frac{\partial f_i(\mu)}{\partial X^T} \) is a constant \( 1 \times N \)-vector (row vector), from the definition of the expectation operator, we have

\[ E[f_k(X)(X_i - \mu_i)] \approx \sigma_i \frac{\partial f_k(\mu)}{\partial X}, \] (2.36a)

\[ E[f_i(X)(X_k - \mu_k)] \approx \frac{\partial f_i(\mu)}{\partial X^T} \sigma_k. \] (2.36b)

Now, using a truncated Taylor series expansion about the mean \( \mu \) of the diffusion functions in Equation (2.34), we obtain:

\[ B_{ik}(X) \approx B_{ik}(\mu) + \frac{\partial B_{ik}(\mu)}{\partial X^T}(X - \mu) + \frac{1}{2}(X - \mu)^T \frac{\partial^2 B_{ik}(\mu)}{\partial X \partial X^T}(X - \mu). \] (2.37)

Taking the expectation, note that by Definition 2.32, each \( B_{ik}(\mu) \in \mathbb{R} \). Since \( E\left[ \frac{\partial B_{ik}(\mu)}{\partial X^T}(X - \mu) \right] = 0 \) (as \( \mu \) is the expected value of \( X \)), we then obtain

\[ E[B_{ik}(X)] \approx B_{ik}(\mu) + E\left[ \frac{1}{2}(X - \mu)^T \frac{\partial^2 B_{ik}(\mu)}{\partial X \partial X^T}(X - \mu) \right]. \]
Further, similar to the derivation of Equation (2.29), the Hessian  
\[ \frac{\partial^2 B_{ik}(\mu)}{\partial X \partial X^T} \] is a symmetric, constant  
\( N \times N \) matrix as the diffusion functions are polynomials  
of at most degree 2. Then by Definition 5.19,

\[
E[B_{ik}(X)] \approx B_{ik}(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\mu)}{\partial X \partial X^T} \sigma \right). \tag{2.38}
\]

Inserting Equations (2.36) and (2.38) back into the expression of (2.34), we  
define the ODEs for the time-evolution of the covariances to be:

\[
\frac{d\sigma_{ik}}{dt} = \frac{\partial f_i(\mu)}{\partial X^T} \sigma \cdot k + \sigma_i \cdot \frac{\partial f_k(\mu)}{\partial X} + \frac{1}{\Omega} \left( B_{ik}(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\mu)}{\partial X \partial X^T} \sigma \right) \right). \tag{2.39}
\]

Then, finally, we have that Equations (2.29) and (2.39) together express the  
full 2MA system:

\[
\frac{d\mu_i}{dt} = f_i(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 f_i(\mu)}{\partial X \partial X^T} \sigma \right)
\]

\[
\frac{d\sigma_{ik}}{dt} = \frac{\partial f_i(\mu)}{\partial X^T} \sigma \cdot k + \sigma_i \cdot \frac{\partial f_k(\mu)}{\partial X} + \frac{1}{\Omega} \left( B_{ik}(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\mu)}{\partial X \partial X^T} \sigma \right) \right)
\]

\[ i, k = 1, ..., N. \]

The Jacobians \( \frac{\partial f_i(\mu)}{\partial X^T} \) and \( \frac{\partial f_k(\mu)}{\partial X} \) of Equation (2.39) account for the noise dynamics of the system and the terms \( \frac{1}{2} \text{tr} \left( \frac{\partial^2 f_i(\mu)}{\partial X \partial X^T} \sigma \right) \) express the individual fluctuations of each event [44,84]. The terms \( \frac{\partial f_i(\mu)}{\partial X^T} \sigma \cdot k \) and \( \sigma_i \cdot \frac{\partial f_k(\mu)}{\partial X} \) indicate that the noise of the pair \((i, k)\) is contributed to in two ways. Let \( l \) be any  
molecule implicated in any reaction involving molecules \( i \) (as a reactant or
a product) and $m$ be any molecule implicated in any reaction with molecule $k$. Then the noise between $i$ and $k$ is given by

1. the relative noise between molecule $i$ and molecules $m$,
2. the relative noise between molecule $k$ and molecules $l$.

Moreover, the weight of the variability between the direct interactions of molecules $i, k$ is further emphasised in the terms $B_{ik}^\mu + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}^\mu}{\partial \mu \partial \mu^T} \sigma \right)$, which express the fluctuation dynamics between these molecules [76].

Interestingly, the formulation of Equation (2.39) mirrors a generalised expression of the Fluctuation-Dissipation Theorem (FDT) of statistical thermodynamics [5, 48, 69].

**Theorem 2.1 (Fluctuation-Dissipation [69]).** Let $\sigma$ be the matrix of covariances, $A$ be the Jacobian matrix of the dynamics for average behaviour and $D$ be a diffusion matrix that depends on the size of the random events. The Fluctuation-Dissipation Theorem of statistical thermodynamics states that the time evolution of covariances is given by

$$\frac{d\sigma}{dt} = A\sigma + \sigma A^T + D.$$  \hspace{1cm} (2.40)

As noted in [69], the FDT as stated above is known under other names depending on the discipline. Indeed, within reaction systems, it is frequently referred to as the Linear Noise Approximation (LNA) [12, 79, 81]. The matrix $A$ of Equation (2.40) is the Jacobian matrix of the drift functions and
the matrix $\mathbf{D}$ can alternatively be viewed as the probability diffusion coefficient matrix which measures the intensity of variations [76]. Per [12], we can write the diffusion matrix $\mathbf{D}$ as $\mathbf{D} = \mathbf{P} \mathbf{P}^T$, for $\mathbf{P} = \mathbf{S} \text{diag}(\mathbf{a}(\mu)) \mathbf{S}^T$.

It should be observed however that there is a difference of terms from Equation (2.39) to Equation (2.40). In fact, the 2MA equations take into account higher order noise than the LNA through the terms $\frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right)$ because of its second-order truncation of the CME. Since we are considering nonlinear reaction rates which introduce random behaviours beyond those described by the LNA/FDT, higher order correlations need to be considered [28, 81].

As a last consideration, we seek to verify that the 2MA equations preserve the symmetry of the covariance matrix. Note that from Equation (2.39), we have that

\[
\frac{d\sigma_{ik}}{dt} = \frac{\partial f_i(\mu)}{\partial \mathbf{x}^T} \sigma_{k} + \sigma_i \frac{\partial f_k(\mu)}{\partial \mathbf{x}} + \frac{1}{\Omega} \left( B_{ik}(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) \right)
\]

\[
\frac{d\sigma_{ki}}{dt} = \frac{\partial f_k(\mu)}{\partial \mathbf{x}^T} \sigma_{i} + \sigma_k \frac{\partial f_i(\mu)}{\partial \mathbf{x}} + \frac{1}{\Omega} \left( B_{ki}(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ki}(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) \right)
\]

(2.41)

Considering first the drift function terms, observe that Equation (2.41) can be rewritten as
\[
\frac{d\sigma_{ki}}{dt} = \left( \sigma_i \frac{\partial f_i(\mu)}{\partial x} \right)^T + \left( \frac{\partial f_i(\mu)}{\partial x^T} \right)\sigma_k + \frac{1}{\Omega} \left( B_{ki}(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ki}(\mu)}{\partial x \partial x^T} \sigma \right) \right).
\]

As \( \sigma_i \frac{\partial f_i(\mu)}{\partial x} \) is scalar, \( \left( \sigma_i \frac{\partial f_i(\mu)}{\partial x} \right)^T = \sigma_i \frac{\partial f_i(\mu)}{\partial x} \) and similarly for \( \frac{\partial f_i(\mu)}{\partial x^T} \sigma_k \), which gives

\[
\frac{d\sigma_{ki}}{dt} = \sigma_i \frac{\partial f_i(\mu)}{\partial x} + \frac{\partial f_i(\mu)}{\partial x^T} \sigma_k + \frac{1}{\Omega} \left( B_{ki}(\mu) + \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ki}(\mu)}{\partial x \partial x^T} \sigma \right) \right).
\]

Considering then the terms involving the diffusion functions, by Equation (2.32) we have

\[
B_{ik}(\mu) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} S_{kj} a_j(\Omega \mu) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{kj} S_{ij} a_j(\Omega \mu) = B_{ki}(\mu).
\]

Furthermore,

\[
\frac{\partial^2 B_{ik}(\mu)}{\partial x \partial x^T} = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} S_{kj} \frac{\partial^2 a_j(\Omega \mu)}{\partial x \partial x^T} = \frac{1}{\Omega} \sum_{j=1}^{M} S_{kj} S_{ij} \frac{\partial^2 a_j(\Omega \mu)}{\partial x \partial x^T} = \frac{\partial^2 B_{ki}(\mu)}{\partial x \partial x^T},
\]

which implies that

\[
\frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\mu)}{\partial x \partial x^T} \sigma \right) = \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ki}(\mu)}{\partial x \partial x^T} \sigma \right).
\]

Accordingly, Equation (2.41) can be expressed as
\[
\frac{d\sigma_{ki}}{dt} = \frac{\partial f_i(\mu)}{\partial x^j} \sigma_{jk} + \sigma_{ij} \frac{\partial f_k(\mu)}{\partial x} + \frac{1}{\Omega} \left( B_{ik}(\mu) + \frac{1}{2} tr \left( \frac{\partial^2 B_{ik}(\mu)}{\partial x \partial x^T} \sigma \right) \right),
\]

and therefore we have that \( \frac{d\sigma_{ik}}{dt} = \frac{d\sigma_{ki}}{dt}, \forall i, k \). Hence, the symmetry of \( \sigma \) is preserved under Equation (2.39) when considered with symmetric initial conditions \( \sigma_{ik}(0) = \sigma_{ki}(0) \).

Having outlined the theoretical basis of the three models in the study of assembly, we now turn to the treatment of their results, as presented in the following section.
Chapter 3

Results

This chapter reports both numerical and mathematical results of the three methods of investigation of the assembly of intermediate filaments (IFs). We begin with results on the chemical master equation (CME) and its exact simulation technique, the stochastic simulation algorithm (SSA). Next, we present mathematical results characterising solutions for both the reaction rate equations (RREs) and the two-moment approximation (2MA) equations. In addition, the asymptotic behaviour of the RREs is characterised. We conclude by making several comparisons: first, to the experimental results of [73], next of all three modelling approaches and finally, of the stochastic framework of the SSA to the deterministic formalism of the RREs to appreciate how the incorporation of the system’s noise influences the behaviour of solutions. Figure 3.1(a) summarises the types of analyses possible for each of the three models. The numerical methods of each of the models are summarised in Figure 3.1(b).
Asymptotic behaviour of solutions

Mathematical tools for small, simple systems only

(a) Analytic Tools

Classic Numerical Techniques for ODEs (Euler methods, Runge-Kutta methods etc.)

(b) Numerical Tools

Figure 3.1: Analytical and numerical tools of the chemical master equation (CME), the two-moment approximation (2MA) equations, and the reaction rate equations (RREs). In (a): The feasibility of analysis of the CME is severely limited, necessitating the use of the 2MA equations (if the noise in the system is of interest), or the RREs (if noise is neglected). In (b): Numerical solutions to the CME are exact in nature as results from the SSA (Stochastic Simulation Algorithm) and the reaction rate equations (RREs). In (a): The feasibility of analysis of the CME is severely limited, necessitating the use of the 2MA equations (if the noise in the system is of interest), or the RREs (if noise is neglected). In (b): Numerical solutions to the CME are exact in nature as results from the SSA (Stochastic Simulation Algorithm) and the reaction rate equations (RREs).
3.1 Chemical Master Equation

We consider now the application of the CME to the polymerisation of IFs. From Section 2.2.2, we have seen how the SSA, based on the properties of the CME, allows for the reconstruction of all possible configurations of the system described by the CME. In order to fully understand the behaviour of the SSA simulation tool, we are interested in characterising its response to several input modifications. This then gives us a fully elucidated view of the system’s inner workings and intricacies.

Simulations of the CME by the SSA produce sample paths which track the change in each population over time. Figure 3.2 shows typical SSA results obtained with one simulation (one numerical experiment), where one path represents the change per molecule over the time. Note that each path exhibits a discrete time jump, that is that a given population will remain constant until it is involved in a reaction $R_j$, at which time it ‘leaps’ to its changed population level.

As originally noted in [21], owing to the random nature of the SSA outlined in Section 2.2.2, each simulation run can produce different results as illustrated in Figure 3.3.
Figure 3.2: **SSA path results**-(a): A sample path from one simulation representing the time evolution of filaments of length 2 and shown in (b) with all populations (for the total number of populations $N = 20$). Each line represents the change in population $N_i$ produced by one SSA simulation. (c): From (b), the length distribution at time $t = 120s$ is reconstructed.
CHAPTER 3. RESULTS

57

(a) Filament of Length 1 \((N_1)\)

(b) Filaments of Length 2 \((N_2)\)

(c) Filaments of Length 3 \((N_3)\)

(d) Average Distribution with Standard Deviations

Figure 3.3: **Twenty simulations of the SSA**—Twenty sample paths generated by the SSA for filaments of (a): length 1, (b): length 2, and (c): length 3. Distributions are averaged and standard deviations are calculated per filament length in (d). This type of result is directly comparable to those of the 2MA equations that govern the dynamics of the means and covariances and is unavailable in the case of the RREs (Section 3.4.2).
In fact, exactly reproducing any given result is extremely rare. Consequently, limited information can be obtained from any one particular SSA simulation [14]. It is for this reason that we calculate average length distributions (the stochastic mean) from numerous simulation runs. This stochastic mean is in fact the most probable configuration of the CME at a given time \( t \) and is also a replicable property. As such it is of much greater interest [14]. The reconstruction of this mean enables us to study the time evolution of distributions, as in Figures 3.3(d) and 3.4. Generally speaking, we observe an increase in the length of filaments produced as time increases, which results in an increase in the mean length (ML; measured in number of ULFs) of the average length distribution, as in Figure 3.4(c).

In Section 2.1 we noted that filament disassembly is not considered in the in vitro system (assembly is an irreversible reaction). Consequently, when experiments are left to react indefinitely, the in vitro IF system will stabilise into one long filament of length \( N \). Once aggregated into one filament, no further reactions are possible. Letting \( N_i \) be the copy number of a filament of length \( i \), we can express this asymptotic aggregation phenomenon mathematically as \((N_1, N_2, ..., N_N) = (0, ..., 0, 1)\) and this result is illustrated in Figure 3.5.

As a final note on the general behaviour of the SSA, consider Table 3.1 which highlights the computational time required for different inputted values. From the reported values within, it is clear that the simulation times were strongly determined by the total number of populations \( N \) which induces changes to the number of potential reactions in the system (Proposition 2.2), and less so by increases to the time of interest \( t \). Higher values of \( N \) also in-
CHAPTER 3. RESULTS

(a) Short Time ($t=60s$)

(b) Long Time ($t=600s$)

(c) Mean Lengths

Figure 3.4: **Length distributions from the SSA**-(a) and (b): SSA distributions in number of ULFs showing the evolution of distributions over increasing time. Additionally, mean lengths are shown to increase as time increases in (c). These results summarise 750 simulations of the SSA.
(a) Increase in Total Number of Populations

<table>
<thead>
<tr>
<th>N</th>
<th>t (s)</th>
<th>Simulation Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>120</td>
<td>0.044543</td>
</tr>
<tr>
<td>100</td>
<td>120</td>
<td>0.134686</td>
</tr>
<tr>
<td>223</td>
<td>120</td>
<td>1.873073</td>
</tr>
<tr>
<td>446</td>
<td>120</td>
<td>18.293660</td>
</tr>
<tr>
<td>892</td>
<td>120</td>
<td>124.642287</td>
</tr>
</tbody>
</table>

(b) Increase in Time

<table>
<thead>
<tr>
<th>N</th>
<th>t (s)</th>
<th>Simulation Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>446</td>
<td>30</td>
<td>10.229234</td>
</tr>
<tr>
<td>446</td>
<td>60</td>
<td>14.851328</td>
</tr>
<tr>
<td>446</td>
<td>120</td>
<td>18.293660</td>
</tr>
<tr>
<td>446</td>
<td>1200</td>
<td>25.114032</td>
</tr>
<tr>
<td>446</td>
<td>2400</td>
<td>26.049322</td>
</tr>
</tbody>
</table>

Table 3.1: **Computational Simulation Time**-The time (in seconds) of one simulation of the SSA for various values of the total number of ULFs $N$ (total number of populations) and six time points. Increases in $t$ do slightly increase the simulation time. However, the computational load is more greatly dependent upon the value of $N$. Simulations were run on an Intel®Core™2 Quad CPU, 2.65 GHz with 8 GB of RAM.

flate the value of $M$, which increases the number of possible reactions. This then implies that the chance of a reaction occurring in the system is more probable, so more events will happen over the time of interest. This ultimately increases the simulation time. To solve this problem, some authors suggest using a $\tau$-leaping scheme, which allows for more than one reaction in any given time interval [7].

Using the average length distributions detailed above, the numerical results presented in this section focus on the simulation of the CME by the SSA. These simulations were carried out to investigate several questions, namely:
CHAPTER 3. RESULTS

1. What is the necessary number of SSA simulation runs to obtain reliable average length distributions (convergence)?

2. How does a change in the total number of populations $N$ affect the average length distribution (effect of changes to the initial condition)?

3. How do variations in reaction rates affect the average length distributions of the SSA (parameter variability)?

To that end, the following standards of comparison were employed. To measure errors in *scalar* values, we used the definition of relative errors given below.

**Definition 3.1.** Let $V_{\text{studied}}$ be the variable of interest and $V_{\text{ref}}$ a reference variable, where the units of $V_{\text{studied}}$ and $V_{\text{ref}}$ are the same. Then the relative error is given by

$$Err_{\text{rel}} = \frac{|V_{\text{studied}} - V_{\text{ref}}|}{V_{\text{ref}}}. \quad (3.1)$$

Two means to compare distributions were employed; the first is a quantification of differences in data sets, and the second a statistical test to measure the equivalency between samples.

**Definition 3.2.** The sum of squared residuals is defined to be

$$SSR = \sum_{i=1}^{N} (Z_i - Y_i)^2,$$

where $Z$ and $Y$ are two samples.
The sum of squared residuals compares two distributions by estimating the
distance between the samples and can therefore be considered as the error
between the reference and studied samples.

Finally, a non parametric Mann-Whitney U-test was used to compare aver-
age length distributions.

**Theorem 3.1.** [54, 75] Suppose there are two samples \((X, Y)\) of size \(m\)
and \(n\), respectively, and let the quantities \(X_1, ..., X_n, Y_1, ..., Y_m\) be arranged
in order. Define the Mann-Whitney U-statistic by

\[
U = \sum_{i=1}^{m} \sum_{j=1}^{n} \tilde{U}_{ij},
\]

where

\[
\tilde{U}_{ij} = \begin{cases} 
0, & X_i > Y_j \\
\frac{1}{2}, & X_i = Y_j \\
1, & X_i < Y_j 
\end{cases}
\]

Then \(U\) can be described as counting the number of times an \(X\) precedes a
\(Y\). We wish to test the null-hypothesis

\[
H_0 : \text{Prob}\{\tilde{U}_{ij} = 1\} = \text{Prob}\{\tilde{U}_{ij} = 0\}
\]

(the samples \(X\) and \(Y\) have identical distributions),

against the alternative hypothesis

\[
H_1 : \text{Prob}\{\tilde{U}_{ij} = 1\} \neq \text{Prob}\{\tilde{U}_{ij} = 0\}
\]
the samples $X$ and $Y$ are not identically distributed).

If $P(U \leq \bar{U}) = \alpha$ under the null-hypothesis, the test will be considered significant on the significance level $\alpha$ if $U \leq \bar{U}$ and the hypothesis of identical distributions of $X$ and $Y$ will be rejected.

By convention, we employed the Mann-Whitney test at a significance level of $\alpha = 0.05$. This implies that, when failing to reject $H_0$, the differences in each of the distributions are statistically insignificant 95% of the time. The probability $P(U > \bar{U})$ is called the $p$-value of the test, and gives the probability that a particular result would occur by chance [38]. Accordingly, smaller $p$-values indicate that extreme results are more likely and hence, we reject the null hypothesis $H_0$ with $p$-values less than $\alpha$.

Mann-Whitney U-tests and $p$-values were calculated using the Matlab® function `ranksum` [56].

Equipped with the measures of comparison outlined above, we begin to investigate the three questions of this section.

### 3.1.1 Necessary Number of Simulations

Owing to the random nature of the SSA, one single simulation run can rarely be reproduced [14]. Accordingly, to accurately characterise the configurations of the system described by the CME at a given time, repeated simulation runs of the SSA are required to construct the average configu-
Figure 3.5: **Asymptotic behaviour**—Filaments assemble into one filament of length $N = 20$ (the total number of ULFs in the system) when left to react over sufficient time $\left((N_1(0), N_2(0), ..., N_{20}(0)) = (20, 0, ..., 0), \ t = 25000s\right)$. 

Figure 3.5: **Asymptotic behaviour**—Filaments assemble into one filament of length $N = 20$ (the total number of ULFs in the system) when left to react over sufficient time $\left((N_1(0), N_2(0), ..., N_{20}(0)) = (20, 0, ..., 0), \ t = 25000s\right)$. 

ration. Define the *convergence* of the SSA to be the number of simulations ($\text{simul}$) giving the most acceptable portrayal of the average length distribution, using a given measure. Here we seek to determine such a $\text{simul}$ to answer our first investigatory question:

1. What is the necessary number of SSA simulation runs to obtain reliable average length distributions (convergence)?

Previous research has established that anywhere from 2000 to $10^{11}$ simulation runs (in the case of very infrequently occurring reactions) are required to be reasonably assured of capturing the true behaviour of the system [19,49]. However, convergence is also affected by ‘the inherent averaging of macroscopic properties of a system of many particles’ [82], so it is worth investi-
gating the appropriate number of SSA runs for any model, as convergence is system size and reaction rate profile dependent. That being said, due to the computational load of the SSA, we seek to balance its convergence with the required simulation time (Table 3.1).

To this end, \( t = 60 \text{s} \) (short time) and \( t = 600 \text{s} \) (long time) were chosen for the study of the necessary number of simulations for tolerable convergence. For both time points, 50, 100, 150, 500, 750, 1000, and 5000 simulations were run. These results were then aggregated to reflect the results of all 7550 simulations. As a first means of comparison, the relative error defined in Definition 3.1 was used with \( V_{\text{studied}} = ML_{\text{simul}} \) and \( V_{\text{ref}} = ML_{\text{ag}} \), where \( ML_x \) is the mean length of the given number of simulations (\( x = \text{simul} \) or \( x = \text{ag} = 7550 \)). Relative error results are reported in Table 3.2. Next, sums of squared residuals were evaluated (Definition 3.2) to measure the differences between average length distributions from each set of simulations and the aggregated result, with findings reported in Table 3.3. Finally, the non-parametric statistical test of Definition 3.1 was applied to the average length distributions of each set of simulations to check for statistical differences in distributions (Table 3.4). Each of the average length distributions, as well as their differences from the aggregated result (per filament length) are shown in Figure 3.6.

All SSA simulations in this section were carried out using rates scaled from the deterministic system according to \( C_j = 1.5 \cdot k_j \cdot 10^{-10} \) (see Section 3.4.1 for a more detailed explanation of this choice).
Table 3.2: **Effect of the Number of Simulations**-Mean lengths (in number of ULFs) over various values of `simul` are compared to the aggregated results (7550 simulations) from which relative errors are calculated. \( N = 446, t = 60 \text{ s} \) (aggregated ML=2.5917 ULFs) and \( t = 600 \) (aggregated ML=6.8362 ULFs). An increase in the number of simulations reduces the relative error of the SSA. An increase in the number of simulations from 5000 to 7550 (the aggregated total) does not produce significant changes to the ML. For example, at \( t = 60 \text{ s} \), this increase would induce only a 0.069653% change in the simulated ML.


Table 3.3: **Effect of the Number of Simulations**—Sums of squared residuals (Definition 3.2) results per filament length from the average length distributions over `simul` simulations as compared to the aggregated average length distribution as reference (`t = 60` and `t = 600`). Since the smallest sum of squared residuals is found with 5000 simulations over both times, 5000 was deemed to be the ideal number of simulations. This is consistent with the results of Table 3.2. Figures 3.6(b) and 3.6(d) show the differences per filament length between distributions over `simul` simulations and the aggregated distribution.

<table>
<thead>
<tr>
<th>simul</th>
<th><code>t = 60s</code></th>
<th><code>t = 600s</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.1959·10^{-4}</td>
<td>84.575·10^{-6}</td>
</tr>
<tr>
<td>100</td>
<td>1.2894·10^{-4}</td>
<td>109.98·10^{-6}</td>
</tr>
<tr>
<td>150</td>
<td>1.5725·10^{-4}</td>
<td>14.897·10^{-6}</td>
</tr>
<tr>
<td>500</td>
<td>0.14611·10^{-4}</td>
<td>8.23893·10^{-6}</td>
</tr>
<tr>
<td>750</td>
<td>0.10421·10^{-3}</td>
<td>1.6134·10^{-6}</td>
</tr>
<tr>
<td>1000</td>
<td>0.28606·10^{-4}</td>
<td>8.4077·10^{-6}</td>
</tr>
<tr>
<td>5000</td>
<td>0.15211·10^{-3}</td>
<td>3.9962·10^{-6}</td>
</tr>
</tbody>
</table>
CHAPTER 3. RESULTS

Figure 3.6: Effect of the Number of Simulations

Higher values of $\text{simul}$ produce smaller sums of squared residuals, as shown in Table 3.3.

(a) $t = 60$ s
(b) $t = 600$ s
(c) $t = 60$ s and $t = 600$ s obtained by the SSA with a different number of simulations ($\text{simul}$). On the right, distributions are subtracted from the aggregated result to show the difference from the mean result for each set.

The left column gives the average final distribution for the aggregated result.
CHAPTER 3. RESULTS

<table>
<thead>
<tr>
<th>simul</th>
<th>p-value</th>
<th>Fail to Reject $H_0$</th>
<th>p-value</th>
<th>Fail to Reject $H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.3580</td>
<td>Yes</td>
<td>0.1812</td>
<td>Yes</td>
</tr>
<tr>
<td>100</td>
<td>0.3589</td>
<td>Yes</td>
<td>0.1750</td>
<td>Yes</td>
</tr>
<tr>
<td>150</td>
<td>0.3589</td>
<td>Yes</td>
<td>0.7030</td>
<td>Yes</td>
</tr>
<tr>
<td>500</td>
<td>0.5964</td>
<td>Yes</td>
<td>0.5657</td>
<td>Yes</td>
</tr>
<tr>
<td>750</td>
<td>0.7293</td>
<td>Yes</td>
<td>0.8507</td>
<td>Yes</td>
</tr>
<tr>
<td>1000</td>
<td>0.7287</td>
<td>Yes</td>
<td>0.5304</td>
<td>Yes</td>
</tr>
<tr>
<td>5000</td>
<td>0.9969</td>
<td>Yes</td>
<td>0.9269</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 3.4: Effect of the Number of Simulations-Mann-Whitney U-test results show average length distributions over all values of simul are statistically equivalent to aggregated results for $t = 60s$ and $t = 600s$.

In all cases, we fail to reject the null hypothesis $H_0$ and can conclude that over both short and long times ($t = 60s$ and $t = 600s$), there is no statistically significant difference between SSA average length distributions over a given number of simulations. Additionally, as an increase in the number of simulations induces an increase in the reported $p$-value, more simulation runs signify a lower chance of rejecting the null hypothesis that distributions are identical.

From both the sums of squared residuals results of Table 3.3 and the $p$-values reported in Table 3.4, it is clear that increasing the number of simulations provides a better convergence. This conclusion is supported by both Table 3.2 and Table 3.3 which reflect the observation that adding 2550 extra simulations to sets of 5000 simulations does not induce a significant change in the results of the SSA. Consequently, we ran 5000 simulations whenever possible. When simulation time was limited, however, 750 simulations were
CHAPTER 3. RESULTS

70

demed to be acceptably convergent and to provide an accurate average of length distributions.

3.1.2 Effect of Initial Condition on SSA Output

In general, we can express the initial condition (IC) as $N(0) = N_0$. For consistency with the experimental work in the IF system, the IC $N(0) = (N, 0, ..., 0)$ was used for all SSA simulations, which in plain language, means that we introduce $N$ ULFs at the beginning of each experiment. Recall that $N$ is also defined to be the total number of populations in the system. Since Table 3.1 demonstrates that the total number of populations can affect the time of simulation, to better understand the behaviour of the SSA, we are interested in investigating how the IC influences the simulation results. This helps to clarify the second question of our study:

2. How does a change in the total number of populations $N$ affect the average length distribution (effect of changes to the initial condition IC)?

The results in this section reflect SSA outputs of 5000 simulations run for $N = 150, t = 1200s$, $N = 300, t = 600s$, and $N = 600, t = 300s$. The length distributions of each set of simulations are juxtaposed in Figure 3.7. To test for the statistical equivalency of each of these length distributions, a Mann-Whitney U-test was undertaken (Definition 3.1) and results of this analysis are reported in Table 3.5.
### Table 3.5: Effect of Initial Condition

Average length distributions in proportions are shown to be statistically equivalent via a Mann-Whitney U-test by comparison of a $k$-fold increase in the initial condition $N$ to a $k$-fold decrease in time.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$t$ (s)</th>
<th>$p$-value</th>
<th>Fail to Reject $H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>600</td>
<td>0.8967</td>
<td>Yes</td>
</tr>
<tr>
<td>600</td>
<td>300</td>
<td>0.8861</td>
<td>Yes</td>
</tr>
<tr>
<td>150</td>
<td>1200</td>
<td>0.8967</td>
<td>Yes</td>
</tr>
<tr>
<td>600</td>
<td>300</td>
<td>0.7823</td>
<td>Yes</td>
</tr>
<tr>
<td>150</td>
<td>1200</td>
<td>0.8861</td>
<td>Yes</td>
</tr>
<tr>
<td>300</td>
<td>600</td>
<td>0.7823</td>
<td>Yes</td>
</tr>
</tbody>
</table>
FIGURE 3.7: Effect of Initial Condition—(a): Distributions in proportions are shown to be statistically equal for a $k$-fold increase in the initial condition $N$ and a $k$-fold decrease in time. (b): The effect of this proportionality is seen when comparing distributions in number of ULFs; the number of filaments of each length $i = 1, \ldots, N$ is approximately doubled with a two-fold increase in the IC and a halving of the time of interest.
From the results detailed herein, we note that when considering average length distributions in proportions, as in Figure 3.7(a), the Mann-Whitney U-test shows them to be identically distributed (Table 3.5) for distributions with a $k$-fold increase in the initial condition $N$ and those with a $k$-fold decrease in time. Moreover, Figure 3.7(b) shows that, when comparing distributions in number of ULFs, a twofold increase in the initial condition $N$ coupled with a halving of the simulation time produces a doubling in the number of filaments of each length. As a result of this proportional equivalency, we could choose to decrease the initial condition (and, therefore, proportionally increase the time) of a given SSA simulation set and remain confident that the results accurately reflect the form of the average length distributions. This is a favourable result in terms of computational load since, as we have seen in Table 3.1, the time of simulation of the SSA of our system is significantly affected by the initial condition $N$.

3.1.3 Effect of Parameter Changes on SSA Output

It is also necessary to grasp the effect the reaction rates have on the system as, with the IC studied in the previous section, they are the only user-inputted parameters used in our simulation of the CME. Thus their influence on the SSA results is believed to be notable. Accordingly, we now concentrate on how the shape of SSA distributions are modified by changes to the reaction rates $C_j$, $j = 1, ..., M$ to answer the final investigatory question of this section:

3. How do variations in reaction rates affect the average length distribu-
tions of the SSA (parameter variability)?

The analyses were carried out with

i) rates scaled from the deterministic system of [73] taken in multiples

ii) random rates of the same order as the rates scaled from the deterministic system

to compare the impact of each set of reaction rates on the shape of the average length distributions. For this, we chose \textit{microscopic} rate parameters in the orders of $10^{-4}$ to $10^{-5}$ for reasons discussed in Section 3.4.1. Random rates were generated using the \textit{rand} function in Matlab, which provides uniformly distributed random numbers (Definition 5.14) on the unit interval $[0, 1]$. It should be noted that the deterministically scaled parameters reflect that all reactions have almost equal probability to occur. However, the random reaction rates exhibit a greater variability and thus include more comparably rare reactions. Figures 3.8(a) and 3.9(a) illustrate the differences in each of the reaction rates under consideration. In Figure 3.8, results obtained from the deterministically scaled rate parameters demonstrate less perceptible differences than those highlighted in Figure 3.9, which compares random rates to the scaled parameters of [73].
and not arbitrary. However, when rate parameters are randomly generated, slight differences appear between distributions. The average length distribution, when varying rates for short and long times, shows the slight differences in the average length distributions with varying rates while (c) and (d) show the slight differences to the average length distributions with varying rates for short and long times. Over short and long times, average length distributions in proportions show less extreme differences than those of Figures 3.9(c) and 3.9(d). Reaction rates are scaled and taken in multiples from the deterministic system of [73].

Figure 3.8: Effect of Parameters

(a) Reaction rates are scaled and taken in multiples from the deterministic system of [73]. (b) shows that MLs change proportionally with the change in rates while (c) and (d) show the slight differences to the average length distributions with varying rates for short and long times. Over short and long times, average length distributions in proportions show less extreme differences than those of Figures 3.9(c) and 3.9(d) when rate parameters are randomly generated.
CHAPTER 3. RESULTS

(a) Random Reaction Rates vs $C_j = k_j \cdot 10^{-10}$ over both short and long times.

(b) Mean Lengths $t = 60$ s, Proportions

(c) $t = 600$ s, Proportions

(d) $t = 600$ s, Proportions

(e) Random rates are compared to rates scaled from the deterministic system. Results from random rates and from scaled parameters are shown to differ significantly.

Figure 3.9: Effect of Parameters - Random rates vs $C_j = k_j \cdot 10^{-10}$ over both short and long times.
CHAPTER 3. RESULTS

Additionally, as investigated in Section 3.1.2 and given the results of Figures 3.8(c) and 3.8(d), we wished to explore whether proportional increases in the parameter coefficients and decreases in the time would produce statistically equivalent distributions. For this study, we compared distributions of \( t = 1200\, s, C_j = k_j \cdot 10^{-10} \) to those from \( t = 600\, s, C_j = 2 \cdot k_j \cdot 10^{-10} \) by again using a Mann-Whitney U-test. We fail to reject the null-hypothesis \( H_0 \) with a \( p \)-value of 0.8103, and conclude that the two are indeed identically distributed. Figure 3.10 illustrates this equivalency. Thus we observe that taking multiples of the reaction rates does nothing but induce faster or slower reactions, thereby affecting the time-scale of the dynamics.

Figure 3.10: Effect of Parameters-Average length distributions in proportions are equivalent when reaction rates are multiplied by a \( k \)-fold increase and are compared to distributions produced over a \( k \)-fold decrease in time \((t = 1200 s, C_j = k_j \cdot 10^{-10} \) and \( t = 600 s, C_j = 2 \cdot k_j \cdot 10^{-10} \)).

Thus we have shown that the shape of the distributions is most definitely impacted by the choice of reaction rate parameters. Varying rates as in Fig-
CHAPTER 3. RESULTS

Figure 3.8(a) can be used to fine-tune the time-scale, and thus produce average length distributions with MLs within ranges of interest. Further, using a Mann-Whitney U-test, we showed a distribution resulting from a $k$-fold increase in the rates is statistically equivalent (in proportion) to one produced with a $k$-fold decrease in time, as seen in Figure 3.10.

In the preceding sections, we answered three pertinent questions on the general behaviour of the SSA studying the in vitro IF system. They can be summarised as follows:

1. Increasing the number of simulations increases the level of accuracy of the average length distributions obtained from the SSA as a representation of the most probable state in the CME system at a given time. Here 5000 simulations were found to be perfectly acceptable, especially considering that increasing the number of simulations by 2550 induced very small improvements in the results. However, in the interest of minimising computational load and time, 750 simulations were decided to produce acceptable average length distributions.

2. Average length distributions taken in proportions with $k$-fold increases in the IC (or total number of populations in the system) are statistically equivalent to distributions resulting from a $k$-fold decrease in the time of interest. Simulations could then be sped up by decreasing $N$ and proportionally increasing the time, as was most expedient in the IF assembly case.

3. The shape of the average length distribution is determined by the type
of reaction rates in the system. Further, taking reaction rates in multiples allows for a fine-tuning of the time-scale of the dynamics. Similar to ICs, a $k$-fold increase in the reaction speed results in statistically identical average length distributions produced with a $k$-fold decrease in time, when considered in proportion.

### 3.2 Reaction Rate Equations

From Equation (2.23), we have the reaction rate equations (RREs) given by

$$
\frac{d\mu_i}{dt} = f_i(\mu), \quad i = 1, \ldots, N, \quad (3.2)
$$

which can also be obtained by keeping only the equations governing the means from the 2MA equations and discounting their stochastic variability. In the following sections, we present results on the nature of solutions to the RREs, as well as the existence and asymptotic behaviour of the system’s equilibria.

#### 3.2.1 Existence, Uniqueness and Positivity of Solutions

Consider the RRE system:

$$
\frac{d\mu_i}{dt} = f_i(\mu) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega \mu), \quad \forall i = 1, \ldots, N, \quad (3.3)
$$

subject to $\mu(0) = \mu_0$.

**Proposition 3.1.** There exists a unique solution to the initial value problem (IVP) given in Equation (3.3).
Proof. Since \( f_i(\mu) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij}a_j(\Omega \mu) \) are all polynomials in \( N \) variables, each \( f_i(\mu) \in C^\infty \). Thus, by Theorem 5.3, there exists a unique solution to the IVP

\[
\frac{d\mu_i}{dt} = f_i(\mu), \quad \mu(0) = \mu_0.
\]

Proposition 3.2. The domain defined as \( D = \{ (\mu_1, \mu_2, ..., \mu_N) | \mu_i \geq 0, \forall i = 1, ..., N \} \) is positively invariant under the flow of Equation (3.3). The solutions of Equation (3.3) are nonnegative within \( D \).

Proof. Consider the point \( \bar{\mu}_i = (\mu_1, \mu_2, ..., 0, ..., \mu_N) \), where the \( i^{th} \) component is 0. Let \( \mathcal{J}_{\text{Reac}} \) be the set of all indices of reactions for which the \( i^{th} \) molecule is a reactant and \( \mathcal{J}_{\text{Prod}} \) be the set of all indices of reactions in which molecule \( i \) is produced. Inserting \( \bar{\mu}_i \) in the corresponding \( \frac{d\mu_i}{dt} \) (i.e. setting \( \mu_i = 0 \) for each respective equation), we obtain:

\[
\frac{d\bar{\mu}_i}{dt} = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij}a_j(\Omega \bar{\mu}_i)
\]

\[
= \frac{1}{\Omega} \sum_{j \in \mathcal{J}_{\text{Reac}}} S_{ij}a_j(\Omega \bar{\mu}_i) + \frac{1}{\Omega} \sum_{j \in \mathcal{J}_{\text{Prod}}} S_{ij}a_j(\Omega \bar{\mu}_i)
\]

\[
= \frac{1}{\Omega} \sum_{j \in \mathcal{J}_{\text{Prod}}} S_{ij}a_j(\Omega \bar{\mu}_i).
\]

The last line follows by the fact that for each of the reactions in which \( i \) is a reactant, \( a_j(\Omega \mu) \) evaluated at \( \bar{\mu}_i = (\mu_1, \mu_2, ..., 0, ..., \mu_N) \) (\( \mu_i = 0 \)), is given
by one of either:

- \( S_{ij}a_j(\Omega \mu_i) = -2C_j \frac{\Omega \mu_i (\Omega \mu_i - 1)}{2} \) (for double reactions),
- \( S_{ij}a_j(\Omega \mu_i) = -C_j \Omega \mu_i \Omega \mu_k \) (for mixed reactions).

Thus, \( a_j(\Omega \mu_i) = 0 \) for all \( j_{Reac} \).

Further, the \( S_{ij} \) will always be positive for \( j_{Prod} \), and as \( \mu_i \geq 0 \) (\( i = 1, \ldots, N \)) and all \( C_j > 0 \), then \( a_j(\Omega \mu_i) \geq 0 \) for \( j \in j_{Prod} \). Hence, by Definition 5.25, the domain \( D \) is positively invariant under the flow of Equation (3.3) as all solutions remain nonnegative, and therefore within \( D \), for nonnegative initial condition \( \mu_0 \). This implies that, solutions to (3.3) are nonnegative (for \( \mu_0 \geq 0 \)), as required.

### 3.2.2 Conservation of Mass

As outlined in Section 2.1, the IF system of consideration is closed, meaning no mass is lost or introduced (for \( t > 0 \)). Proposition 3.3 translates this physical phenomena into an equation for the conservation of mass within the RREs. Further, we show in Proposition 3.5 that this conservation equation can be used to demonstrate the global stability of the unique equilibrium \((0,0,0,...,\frac{1}{\Omega})\) when it is used as a Lyapunov function.

**Proposition 3.3.** For the IF system, there is conservation of mass for the RREs given in Equation (3.3) which can be expressed as:

\[
\sum_{i=1}^{N} i \mu_i(t) = \frac{N}{\Omega},
\]
CHAPTER 3. RESULTS

where \( \frac{N}{\Omega} \) is the initial concentration in ULFs.

Proof. We seek to show that:

\[
\sum_{i=1}^{N} i \frac{d\mu_i}{dt} = 0 \iff \sum_{i=1}^{N} i\mu_i(t) = k, \ \forall t \geq 0,
\]

by integration over \( t \), and for some \( k \in \mathbb{R}_+ \), as solutions to (3.3) are non-negative by Proposition 3.2. To begin, consider the sum:

\[
\sum_{i=1}^{N} i \frac{d\mu_i}{dt} = \sum_{i=1}^{N} \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega \mu),
\]

where \( f_i(x) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega x) \) for \( x \) given in concentration (Equation (2.22)). As noted in Section 2.1, there exist two classes of reactions in the model of IF assembly: double reactions, involving but one filament length as a reactant, and mixed reactions, involving the reaction of two different filament lengths. Let \( J \) be the set of indices of all reactions. Interchanging summation, and partitioning \( J \) into two subsets: indices of double reactions \( J_{doub} \) and indices of mixed reactions \( J_{mix} \), Equation (3.5) becomes:

\[
\sum_{i=1}^{N} i \frac{d\mu_i}{dt} = \frac{1}{\Omega} \left( \sum_{j \in J_{doub}} \sum_{i=1}^{N} i S_{ij} a_j(\Omega \mu) + \sum_{j \in J_{mix}} \sum_{i=1}^{N} i S_{ij} a_j(\Omega \mu) \right).
\]

Then, considering first the reactions \( j \) in the set \( J_{doub} \), the \( S_{ij} a_j(\Omega \mu) \) can take two possible forms:
\[ S_{ij}a_j(\Omega \mu) = \begin{cases} 
-2C_j \frac{\Omega \mu_i(\Omega \mu_i - 1)}{2}, & \text{for reactants of reaction } j \in J_{doub}, \\
C_j \frac{\Omega \mu_i(\Omega \mu_i - 1)}{2}, & \text{for products of reaction } j \in J_{doub}. 
\end{cases} \]

Inserting these expressions into \( A \) of the RHS of Equation (3.6) we obtain:

\[
A = \sum_{j \in J_{doub}} \sum_{i=1}^{\lfloor \frac{N}{2} \rfloor} \left( -iC_j \Omega \mu_i(\Omega \mu_i - 1) + 2iC_j \Omega \mu_i(\Omega \mu_i - 1) \right) = 0. \quad (3.7)
\]

Note that the second term in the above sum follows from the fact that any double reaction produces a filament of twice the length of its reactant molecules and the bounds follow from Proposition 2.1 (as only the first \( \lfloor \frac{N}{2} \rfloor \) lengths will react in double reactions). Similarly, considering reactions \( j \in J_{mix} \) in which \( \mu_i \) and \( \mu_k \) assemble, we have:

\[ S_{ij}a_j(\Omega \mu) = \begin{cases} 
-C_j \Omega \mu_i \Omega \mu_k, & \text{for reactants of reaction } j \in J_{mix}, \\
C_j \Omega \mu_i \Omega \mu_k, & \text{for products of reaction } j \in J_{mix}. 
\end{cases} \]

Accordingly, let filaments of length \( i \) and \( k \) react in the mixed reaction \( j \), to give a filament of length \( i + k \). The term \( B \) of Equation (3.6) become:
\[ B = \sum_{j \in J_{mix}} \sum_{i=1}^{N-i} \sum_{k=i+1}^{N-i} \left( i(-C_j \Omega \mu_i \Omega \mu_k) + k(-C_j \Omega \mu_i \Omega \mu_k) \right) + (i + k)(C_j \Omega \mu_i \Omega \mu_k) = 0, \quad (3.8) \]

with bounds given by Proposition 2.1 (as, by convention, we only consider lengths \( i < k \) in mixed reactions and can form, at most, a filament of length \( N \)). Thus combining both (3.7) and (3.8), Equation (3.6)

\[ \sum_{i=1}^{N} i \frac{d\mu_i}{dt} = 0 \]

holds. Integrating with respect to \( t \) yields:

\[ \sum_{i=1}^{N} i \mu_i(t) = k, \quad t \geq 0, \]

for some \( k \), which can be uniquely determined by the initial condition (IC).

Let \( t = 0 \). By the above sum we have that

\[ k = \sum_{i=1}^{N} i \mu_i(0), \]

and given, for example, the IC reflecting the experimental conditions of IF assembly \( \mu_1(0) = \frac{N}{\Omega}, \mu_i(0) = 0 \) (for \( i = 2, ..., N \)), we can express \( k \) as:

\[ k = \sum_{i=1}^{N} i \mu_i(0) = 1 \cdot (\mu_1(0)) = \frac{N}{\Omega}. \]
CHAPTER 3. RESULTS

Thus,
\[ \sum_{i=1}^{N} i\mu_i(t) = \frac{N}{\Omega}, \quad \forall t \geq 0, \]
and there is conservation of mass for the IF system in the case of the RREs.

3.2.3 Existence of Unique Equilibrium Point

**Proposition 3.4.** For the IF system, the RREs given in Equation (3.3) have the unique, positive equilibrium \( \hat{\mu} = (0, \ldots, 0, \frac{1}{\Omega}) \).

**Proof.** Considering the general expression of the RREs:
\[ \frac{d\mu_i}{dt} = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega \mu), \quad i = 1, \ldots, N \]
and per Definition 5.26, the equilibria of the RRE system are found when:
\[ 0 = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega \mu), \]
for \( i = 1, \ldots, N \). In particular, the equation governing \( \frac{d\mu_N}{dt} \) is:
\[ \frac{d\mu_N}{dt} = C_{N-1}\mu_1\mu_{N-1} + \ldots + C_M\mu_a\mu_b, \]
where \( a, b = \left\lfloor \frac{N}{2} \right\rfloor = \frac{N}{2} \) (if \( N \) is even) or \( a = \left\lfloor \frac{N}{2} \right\rfloor = \frac{N-1}{2} \), \( b = N - a \) (if \( N \) is odd), as outlined in Proposition 2.1. Thus, at equilibrium, we require:
\[ 0 = C_{N-1}\mu_1\mu_{N-1} + \ldots + C_M\mu_a\mu_b. \tag{3.9} \]
Since $C_j > 0$, $(j = 1, ..., M)$ and by Proposition 3.2, we have the nonnegativity of solutions $(\mu_i \geq 0$, $i = 1, ..., N)$, Equation (3.9) is a sum of nonnegative values. Thus the only solution satisfying (3.9) is $\mu_i = 0$, $i = 1, ..., N - 1$. Therefore as all $a_j(\Omega \mu)$ are polynomials of degree 2 with no zero order terms and $\mu_N$ is not a reactant in any reaction, $a_j(\Omega \tilde{\mu}) = 0$, $\forall j$. This implies that $\frac{d\mu_i}{dt} = 0$ for $i = 1, ..., N$ with $\tilde{\mu} = (0, ..., 0, \mu_N)$. Now, by the conservation of mass,

$$\frac{N}{\Omega} = \sum_{i=1}^{N} i\tilde{\mu}_i(t) = N\mu_N \iff \mu_N = \frac{1}{\Omega}.$$ 

Then, since $a_j(\Omega \tilde{\mu}) = a_j(0, ..., 0, 1) = 0$ for $j = 1, .., M$, and

$$\frac{d\mu_i}{dt} = f_i(\tilde{\mu}) = 0, \quad i = 1, ..., N,$$

the RREs for the IF system have a unique equilibrium point $(0, ..., 0, \frac{1}{\Omega})$. □

### 3.2.4 Global Stability Analysis

**Proposition 3.5.** For the RREs given in Equation (3.3) the unique, positive equilibrium point $(0, ..., 0, \frac{1}{\Omega})$ is globally asymptotically stable (GAS).

**Proof.** By Equation (3.4), we can reduce the system in Equation (3.3) to being (N-1)-dimensional:
\[ \mu_N(t) = \frac{1}{\Omega} - \frac{1}{N} \sum_{i=1}^{N-1} i\mu_i(t), \]
\[ \frac{d\mu_i}{dt} = f_i(\mu) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega\mu), \quad (i = 1, ..., N - 1). \quad (3.10) \]

Let \( \tilde{D} = \{(\mu_1, \mu_2, ..., \mu_{N-1}) | \mu_i \geq 0, \forall i = 1, ..., N - 1 \} \) be positively invariant under the flow of the reduced system (as proven for \( D \) under the flow of the \( N \)-dimensional system in Proposition 3.2). From Proposition 3.4, \( \tilde{\mu} = (0, 0, ..., 0) \) is the unique equilibrium of the system of Equation (3.10).

Using the conservation of mass equation, we construct a Lyapunov function (Definition 5.27) to prove the asymptotic stability of the zero equilibrium \( \tilde{\mu} \) of the reduced system (3.10).

Consider the Lyapunov function \( V(\mu) = \sum_{i=1}^{N-1} i\mu_i(t) \) on \( \tilde{D} \). Then \( V(0) = 0 \), and \( V(\mu) > 0 \) for \( \mu \in \tilde{D}\setminus\{0\} \). Therefore by Definition 5.27, \( V(\mu) \) is positive definite on \( \tilde{D} \).

Let \( J^N \) be the set of indices of any reaction \( R_j \) creating a filament of length \( N \), and \( J_{doub}^{N-1} \) and \( J_{mix}^{N-1} \) be the sets of all double and mixed reaction indices which do not create a filament of length \( N \). Differentiating the Lyapunov function \( V(\mu) \), we obtain:
\[
\frac{dV(\mu)}{dt} = \sum_{i=1}^{N-1} \frac{d\mu_i}{dt} = \sum_{i=1}^{N-1} \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} a_j(\Omega \mu)
\]

\[
= \frac{1}{\Omega} \left( \sum_{j \in J_{doub}}^{N-1} \sum_{i=1}^{1} iS_{ij} a_j(\Omega \mu) + \sum_{j \in J_{mix}}^{N-1} \sum_{i=1}^{1} iS_{ij} a_j(\Omega \mu) + \sum_{n \in J}^{N-1} \sum_{i=1}^{1} iS_{in} a_n(\Omega \mu) \right).
\]

Using the same argument as in Equations (3.7) and (3.8) but with \(i = 1, ..., \left\lfloor \frac{N-1}{2} \right\rfloor\), \(A = 0\) and \(B = 0\), giving

\[
\frac{dV(\mu)}{dt} = \frac{1}{\Omega} \sum_{n \in J}^{N-1} \sum_{i=1}^{1} iS_{in} a_n(\Omega \mu). \tag{3.11}
\]

Note that \(a_n(\Omega \mu)\) has two possible forms:

1. \(C_n \frac{\Omega \mu_i (\Omega \mu_i - 1)}{2}, \quad \forall i \in \{m | 2m = N\}\) (double reaction),

2. \(C_n \Omega \mu_i \Omega \mu_k, \quad \forall (i, k) \in \{(m, p) | m + p = N\}\) (mixed reaction).

In both cases, for \(\mu \in \tilde{D} \setminus \{0\}\), Equation (3.11) has propensity functions \(a_n(\Omega \mu) > 0\). Further, for all reactions \((R_n, n \in J_N)\) producing a filament of length \(N\), \(i\) is only ever a reactant. Accordingly

1. \(S_{in} = -2\) if \(n \in J_N\) is a double reaction,

2. \(S_{in} = -1\) if \(n \in J_N\) is a mixed reaction,
CHAPTER 3. RESULTS

and hence $S_{in} < 0$, $\forall i$ in Equation (3.11). Therefore $\frac{dV(\mu)}{dt} < 0$ for $\mu \in \tilde{D}\{0\}$ and by Theorem 5.4, the unique zero-equilibrium of the reduced (N-1)-dimensional system of Equation (3.10) is asymptotically stable with respect to $\tilde{D}$. Furthermore, we have that

$$\lim_{t \to +\infty} \mu_N(t) = \lim_{t \to +\infty} \left( \frac{1}{\Omega} - \frac{1}{N} \sum_{i=1}^{N-1} i\mu_i(t) \right) = \frac{1}{\Omega},$$

which implies the unique equilibrium point $(0, 0, ..., 0, \frac{1}{\Omega})$ of the N-dimensional system is asymptotically stable with respect to the domain D, and since D represents the entire domain of interest, $(0, 0, ..., 0, \frac{1}{\Omega})$ is globally asymptotically stable with respect to D.

The solution, equilibrium and asymptotic behaviour identified within this section reflect those of the mass-action description of the system. This is despite the slight differences in the RRE description of double reactions (as discussed at the end of Section 2.3), which illustrates the analogous nature of the RREs and the Smoluschowski coagulation equation [29].

3.3 Two Moment Approximation

Finally, as has been noted throughout, accounting for stochasticity in assembly models by means of the CME makes mathematical analyses generally untractable as a result of the uncountable state space on which it is defined. Since the 2MA equations represent a moment truncation of the CME, their reduced dimensionality makes obtaining mathematical results from the 2MA equations more feasible. This section concerns the characterisations of so-
olutions to the 2MA equations and presents the existence of one equilibrium in addition to mass conservation results.

### 3.3.1 Existence and Uniqueness of Solutions

Recall the two-moment approximation (2MA) equations of Section 2.4 are:

\[
\frac{d\mu_i}{dt} = f_i(\mu) + \frac{1}{2} tr \left( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \sigma \right) \tag{3.12a}
\]

\[
\frac{d\sigma_{ik}}{dt} = \frac{\partial f_i(\mu)}{\partial x^T} \sigma_k + \sigma_i \frac{\partial f_k(\mu)}{\partial x} + \frac{1}{\Omega} \left( B_{ik}(\mu) + \frac{1}{2} tr \left( \frac{\partial^2 B_{ik}(\mu)}{\partial x \partial x^T} \sigma \right) \right) \tag{3.12b}
\]

\[i, k = 1, \ldots, N\]

subject to the initial condition (IC) \(\mu(0) = \mu_0, \sigma(0) = \sigma_0\).

**Proposition 3.6.** There exists a unique solution to the initial value problem (IVP) defined in Equations (3.12).

**Proof.** By Equations (2.22) and (2.32), the drift and diffusion functions \(f_i(\mu)\) and \(B_{ik}(\mu)\), respectively, are linear combinations of the propensity functions \(a_j(\Omega \mu)\) given in Definition 2.3. Since each \(a_j(\Omega \mu)\) is a polynomial, \(f_i(\mu), B_{ik}(\mu)\) are also polynomials, so they both have continuous derivatives of all orders (i.e. \(f(\mu), B_{ik}(\mu) \in C^\infty\)). Hence, by Theorem 5.3, there exists a unique solution to the IVP defined in Equations (3.12). \(\square\)
3.3.2 Equilibrium Results

We have seen in Section 3.1 that left to react indefinitely, the SSA results stabilise to one filament of length $N$, or in concentrations, to $(X_1, ..., X_{N-1}, X_N) = (0, ..., 0, \frac{1}{\Omega})$. Not only does this reflect the experimental observations of [73], but it was shown in Section 3.2.3 that the reaction rate equations (RREs) have the analogous unique, positive equilibrium

$$\tilde{\mu} = (\mu_1, ..., \mu_{N-1}, \mu_N) = \left(0, ..., 0, \frac{1}{\Omega}\right). \tag{3.13}$$

Considering $\tilde{\mu}$ as above, note that the drift and diffusion functions $f_i$ and $B_{ik}$ depend only on $\mu$. Since $\tilde{\mu}$ is an equilibrium of the RRE system, we have that $f_i(\tilde{\mu}) = 0$, $\forall i$. Then the drift function terms in Equation (3.12a) evaluated at $\tilde{\mu}$ are also equal to 0.

Further, owing to the nature of the reactions in the IF system given as in (2.1) and (2.2), all $a_j(\Omega \mu)$ are polynomials in one or two variables of at most degree 2 with no zeroth order degree. Given that the diffusion functions $B_{ik}(x)$ ($x$ in concentration) are expressed as $B_{ik}(x) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} S_{kj} a_j(\Omega x)$, and since $\tilde{\mu}_i = 0$, $i = 1, ..., N - 1$ and $\mu_N$ is never a reactant of any $R_j$, $B_{ik}(\tilde{\mu}) = \frac{1}{\Omega} \sum_{j=1}^{M} S_{ij} S_{kj} a_j(\Omega \tilde{\mu}) = 0$. As a result, the diffusion functions $B_{ik}(\mu)$ of Equation (3.12b) evaluated at $\tilde{\mu}$ are equal to 0.

Accordingly, let $\tilde{\sigma}$ be the component of an equilibrium point for the covariances and we can then express the equilibria of the 2MA equations (3.12) as satisfying the following equations
$0 = \frac{1}{2} \text{tr} \left( \frac{\partial^2 f_i(\tilde{\mu})}{\partial x \partial x^T} \tilde{\sigma} \right)$ \hspace{1cm} (3.14a)

$0 = \frac{\partial f_i(\tilde{\mu})}{\partial x^T} \tilde{\sigma} + \frac{\partial f_k(\tilde{\mu})}{\partial x} + \frac{1}{\Omega} \left( \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\tilde{\mu})}{\partial x \partial x^T} \tilde{\sigma} \right) \right)$. \hspace{1cm} (3.14b)

These equations are used to prove the next proposition.

**Proposition 3.7.** The point in $\mathbb{R}^{N+N^2}$

$$(\tilde{\mu}, \tilde{\sigma}) = (\mu_1, \ldots, \mu_{N-1}, \mu_N, \sigma_{11}, \ldots, \sigma_{n-1n}, \sigma_{nn})$$

$$= (0, \ldots, 0, \frac{1}{\Omega}, 0, \ldots, 0)$$

is an equilibrium point of the 2MA equations given in Equation (3.12).

**Proof.** Let $\tilde{\sigma} = (0, 0, \ldots, 0)$, which can also be expressed as the $N \times N$ zero matrix $0$ and consider (3.14):

$$0 = \frac{1}{2} \text{tr} \left( \frac{\partial^2 f_i(\tilde{\mu})}{\partial x \partial x^T} \tilde{\sigma} \right) = \frac{1}{2} \text{tr} \left( \frac{\partial^2 f_i(\tilde{\mu})}{\partial x \partial x^T} 0 \right) = 0,$$

$$0 = \frac{\partial f_i(\tilde{\mu})}{\partial x^T} 0 + \frac{\partial f_k(\tilde{\mu})}{\partial x} + \frac{1}{\Omega} \left( \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\tilde{\mu})}{\partial x \partial x^T} 0 \right) \right)$$

$$= \frac{\partial f_i(\tilde{\mu})}{\partial x^T} 0 + \frac{\partial f_k(\tilde{\mu})}{\partial x} + \frac{1}{\Omega} \left( \frac{1}{2} \text{tr} \left( \frac{\partial^2 B_{ik}(\tilde{\mu})}{\partial x \partial x^T} 0 \right) \right) = 0$$

confirming that $$(\tilde{\mu}, \tilde{\sigma}) = (\mu_1, \ldots, \mu_{N-1}, \mu_N, \sigma_{11}, \ldots, \sigma_{n-1n}, \sigma_{nn}) = (0, \ldots, 0, \frac{1}{\Omega}, 0, \ldots, 0)$$ is an equilibrium point of the 2MA equations (3.12). \qed
3.3.3 Conservation of Mass and Notes on the 2MA Equations

Considering only the dynamics of the means of Equation (3.12a) given by

$$\frac{d\mu_i}{dt} = f_i(\mu) + \frac{1}{2} tr \left( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \sigma \right), \forall i = 1, \ldots, N$$

we show there is conservation of mass for the 2MA equations.

**Proposition 3.8.** For the 2MA equations of the IF system, there is conservation of mass for the dynamics governing the means of Equation (3.15) according to:

$$\sum_{i=1}^{N} i \mu_i(t) = k,$$

where $k \in \mathbb{R}$.

**Proof.** Before presenting the main result, we construct a linear system of equations for the $N$ traces $tr \left( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \sigma \right)$ in Equation (3.15).

Begin by observing that we can express the reaction index $j$ as $j = k + \sum_{n=1}^{i-1} (N - 2n)$ for $i = 1, \ldots, \left\lfloor \frac{N}{2} \right\rfloor, k = i, \ldots, N - i$, with bounds as given in Proposition 2.1. Considering this expression for $j$, we make the following observations.

1. The numbers of reactions involving $i$ as the smallest reactant is $\beta_i = N - 2i + 1$, by Proposition 2.1. From the expression for $j$ given above, to calculate the number of reactions involving $i$ as the smallest reactant,
fix $i$ and count $k$ such as $\sum_{k=i}^{N-i} 1 = N - i - i + 1 = \beta_i$. Hence $j$ as defined above is consistent with Proposition 2.1.

2. Let $k = i$,

$$j = i + \sum_{n=1}^{i-1} (N - 2n) = i + N(i - 1) - \frac{2(i - 1)i}{2}.$$  

Then $j$ is the index of the first reaction involving $i$. By Proposition 2.1, $j$ is also given by the sum of the number of reactions involving all lengths 1 to $i - 1$ (all previous smallest reactants) plus 1 (for the index of the current reaction), as follows

$$\left(\sum_{n=1}^{i-1} \beta_n\right) + 1 = \left(\sum_{n=1}^{i-1} N - 2n + 1\right) + 1 = i + N(i - 1) - \frac{2(i - 1)i}{2}.$$  

3. Finally, let $i = 1$ and $k = 1$. The index of the first reaction is $j = 1 + \sum_{n=1}^{0} (N - 2n) = 1$. Letting $i = \left\lfloor \frac{N}{2} \right\rfloor$ and $k = N - \left\lfloor \frac{N}{2} \right\rfloor$, the index of the last reaction is given by

$$j = N - \left\lfloor \frac{N}{2} \right\rfloor + \sum_{n=1}^{\left\lfloor \frac{N}{2} \right\rfloor - 1} (N - 2n) = N - \left\lfloor \frac{N}{2} \right\rfloor + N \left(\left\lfloor \frac{N}{2} \right\rfloor - 1\right)$$  $$- 2 \left(\left\lfloor \frac{N}{2} \right\rfloor - 1\right) \frac{\left\lfloor \frac{N}{2} \right\rfloor}{2}$$  $$= N \left[\left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{N}{2} \right\rfloor^2\right],$$
which, from Proposition 2.2 is equivalent to

\[ M = \sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} (N - 2i + 1) = N \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{N}{2} \right\rfloor^2. \]

Thus \( j = 1, \ldots, M \).

We can then view every pair \((i,k)\) for \( i = 1, \ldots, \left\lfloor \frac{N}{2} \right\rfloor, k = i, \ldots, N - i \) as the indices of the reactants of reaction \( R_j, j = k + \sum_{n=1}^{i-1}(N - 2n), \) with \( j = 1, \ldots, M \). Further, note that the second observation provides a novel way to count the jumps from one horizontal line to another in the transposed stochiometric matrix \( S^T \) defined in Definition 2.4.

Consider then the Hessian of each drift function \( f_i(\mu) \) given by \( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \) with elements \( h_{mn} \). By definition, the trace \( tr \left( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \sigma \right) \) is given by:

\[ h_{11}\sigma_{11} + h_{12}\sigma_{21} + \ldots + h_{1N}\sigma_{N1} + h_{21}\sigma_{12} + \ldots + h_{2N}\sigma_{N2} + \ldots + h_{NN}\sigma_{NN} \quad (3.17) \]

which is a linear function of the \( N^2 \) variables \( \sigma_{ik} \). Since each drift function given in Equation (2.22) is a polynomial of degree 2, each \( \frac{\partial^2 f_i(\mu)}{\partial x \partial x^T} \) is a constant, symmetric \( N \times N \) matrix. This then implies that \( h_{nm} = h_{mn} \) for all \( m, n \). We can thus rewrite Equation (3.17) to be
Additionally by the symmetry of the covariance matrix $\sigma$, we have the equality $\sigma_{nm} = \sigma_{mn}$. As such, Equation (3.17) can be further reduced to be expressed as:

\[
    h_{11}\sigma_{11} + h_{12}\sigma_{21} + ... + h_{1N}\sigma_{N1} + h_{12}\sigma_{12} + ... + h_{2N}\sigma_{N2} + ... h_{NN}\sigma_{NN}.
\]

or in condensed form as

\[
    \text{tr}\left( \frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) = \sum_{m=1}^{N} h_{mm}\sigma_{mm} + \sum_{m=1}^{N} \sum_{n=m+1}^{N-m} 2h_{mn}\sigma_{mn}.
\]

Note that the coefficients $h_{mn}$ necessarily depend on the molecule $i$ of consideration in $\text{tr}\left( \frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right)$. By the definition of each $f_i(\mu)$ as the sums of the propensity functions of reactions involving $i$, let $J_{\text{doub}}$ be the set of all indices of double reactions and $J_{\text{mix}}$ the set of all indices of mixed reactions. Considering double reactions, we have that the entries of each Hessian $\frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T}$ are given by

- $S_{ij}C_j = -2C_j$ when $i$ is a reactant in $R_j$, $j \in J_{\text{doub}}$

- $S_{ij}C_j = C_j$ when $i$ is a product of reaction $R_j$, $j \in J_{\text{doub}}$.

Similarly, for mixed reactions, the elements of each $\frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T}$ are given by
\[ S_{ij}C_j = -C_j \] when \( i \) is a reactant in \( R_j, \ j \in J_{mix} \)

\[ S_{ij}C_j = C_j \] when \( i \) is a product of reaction \( R_j, \ j \in J_{mix} \).

All other elements are 0. Moreover, considering the double and mixed reactions of the system, Equation (3.19) becomes:

\[
tr \left( \frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) = \left\lfloor \frac{N}{2} \right\rfloor \sum_{m=1}^{\left\lfloor \frac{N}{2} \right\rfloor} h_{mm} \sigma_{mm} + \sum_{m=1}^{\left\lfloor \frac{N}{2} \right\rfloor} \sum_{n=m+1}^{N-m} h_{mn} \sigma_{mn}. \tag{3.19}
\]

The first terms of Equation (3.19) are given as such since by Proposition 2.1, only the first \( m = 1, \ldots, \left\lfloor \frac{N}{2} \right\rfloor \) molecules are reactants in double reactions. The mixed reactions in the system are accounted for in the second terms of Equation (3.19). Thus, though the specific coefficient \( h_{mm} \) depends on the molecule \( i \) of consideration in each trace, the variables \( \sigma_{mn} \) are independent of this \( i \).

We count the number of the variables \( \sigma_{mn} \):

\[
\sum_{m=1}^{\left\lfloor \frac{N}{2} \right\rfloor} 1 + \sum_{m=1}^{\left\lfloor \frac{N}{2} \right\rfloor} \sum_{n=m+1}^{N-m} 1 = \sum_{m=1}^{\left\lfloor \frac{N}{2} \right\rfloor} (1 + N - m - (m + 1) + 1) = \sum_{m=1}^{\left\lfloor \frac{N}{2} \right\rfloor} (N - 2m + 1) = M,
\]

meaning each trace \( tr \left( \frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) \) is a linear function of at most the same \( M \) variables \( \sigma_{mn} \) (since \( h_{mn} \) can be 0). Then, since there are then \( N \) traces \( tr \left( \frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) \)-one for each \( \frac{du}{dt} \)-letting \( \varsigma_j = \sigma_{ik} \), for \( j = k + \sum_{n=1}^{i-1} N - 2n \), we can express these \( N \) linear functions of the \( M \) variables as the system...
$F_\varsigma$ given by

$$
F_\varsigma = \begin{pmatrix}
-2C_1 & -2C_2 & -2C_3 & -2C_4 & \cdots & 0 & 0 & \cdots & 0 \\
C_1 & -2C_2 & 0 & 0 & \cdots & -2CN & -2C_{N+1} & \cdots & 0 \\
0 & 2C_2 & -2C_3 & 0 & \cdots & 0 & -2C_{N+1} & \cdots & 0 \\
0 & 0 & 2C_3 & -2C_4 & \cdots & C_N & 0 & \cdots & 0 \\
0 & 0 & 0 & 2C_4 & \cdots & 0 & 2C_{N+1} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
-2C_{N} & -2C_{N+1} & \cdots & 0 & \cdots & 0 & -2C_1 & -2C_2 & -2C_3 \\
\end{pmatrix}
$$

(3.20)

Note that each element $F_{ij}$ indicates whether molecule $i$ is

- not a reactant or product of $R_j$ ($F_{ij} = 0$),
- a reactant of a double reaction $R_j$ ($F_{ij} = S_{ij}C_j = -2C_j$),
- a reactant of a mixed reaction $R_j$ ($F_{ij} = 2S_{ij}C_j = -2C_j$),
- a product of a double reaction $R_j$ ($F_{ij} = S_{ij}C_j = C_j$),
- a product of a mixed reaction $R_j$ ($F_{ij} = 2S_{ij}C_j = 2C_j$),

where the factors of 2 in the case of mixed reactions are attributed the symmetry of $\sigma$, as in Equation (3.18). The structure of $F$ is the same as that of the stochiometric matrix $S$ of Definition 2.4. This equivalency of form is given naturally as the $i^{th}$ rows of both $F$ and $S$ account for all reactions implicating the $i^{th}$ molecule. Accordingly, for each reaction $j = 1, \ldots, M$, they both have non-zero entries in only reactant and product positions.

Then considering the system of Equation (3.20), we examine the conservation of the mass of Equation (3.15) as follows. We seek to show that
\[
\sum_{i=1}^{N} i \frac{d\mu_i}{dt} = 0 \iff \sum_{i=1}^{N} i \mu_i(t) = k, \ \forall t \geq 0,
\]

by integration over \( t \) and for some \( k \in \mathbb{R} \). Since

\[
\sum_{i=1}^{N} i \frac{d\mu_i}{dt} = \left( \sum_{i=1}^{N} i f_i(\mu) + \sum_{i=1}^{N} \frac{1}{2} tr \left( \frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) \right), \quad (3.21)
\]

from Proposition 3.3, we have that \( A = 0 \). Considering then the term \( B \), recall the system of Equation (3.20), with \( \mathbf{F} \) the constant \( N \times M \) matrix defined therein. We can then rewrite \( B \) of Equation (3.21) to be:

\[
B = \frac{1}{2} \sum_{i=1}^{N} i tr \left( \frac{\partial^2 f_i(\mu)}{\partial \mathbf{x} \partial \mathbf{x}^T} \sigma \right) = \frac{1}{2} \sum_{i=1}^{N} i F_i \cdot \varsigma, \quad (3.22)
\]

where \( F_i \) denotes the \( i \)th row of \( \mathbf{F} \). Recall that the entries of \( \mathbf{F} \) can be summarised as

\begin{itemize}
  \item \( F_{ij} = 0 \) (\( i \) not a reactant or product of \( R_j \)),
  \item \( F_{ij} = -2C_j \) (\( i \) a reactant of any reaction \( R_j \)),
  \item \( F_{ij} = C_j \) (\( i \) a product of a double reaction \( R_j \)),
  \item \( F_{ij} = 2C_j \) (a product of a mixed reaction \( R_j \)).
\end{itemize}

Accordingly, let \( J_{\text{doub}} \) be the set of all indices of double reactions and \( J_{\text{mix}} \) be the set of all indices of mixed reactions. We can rewrite Equation (3.22) to be

\begin{itemize}
  \item \( F_{ij} = 0 \) (\( i \) not a reactant or product of \( R_j \)),
  \item \( F_{ij} = -2C_j \) (\( i \) a reactant of any reaction \( R_j \)),
  \item \( F_{ij} = C_j \) (\( i \) a product of a double reaction \( R_j \)),
  \item \( F_{ij} = 2C_j \) (a product of a mixed reaction \( R_j \)).
\end{itemize}
\[ B = \frac{1}{2} \left( \sum_{j \in J_{\text{doub}}} \sum_{i=1}^{N} iF_{ij} \varsigma_j + \sum_{j \in J_{\text{mix}}} \sum_{i=1}^{N} iF_{ij} \varsigma_j \right). \]  

(3.23)

Considering first term \( D \), we have that

\[ iF_{ij} = \begin{cases} 
-2iC_j, & \text{when } i \text{ is a reactant in the double reaction } R_j, \\
iC_j, & \text{when } i \text{ is a product in the double reaction } R_j.
\end{cases} \]

So \( D \) becomes

\[ D = \left( \sum_{j \in J_{\text{doub}}} \sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} -2iC_j \varsigma_j + 2iC_j \varsigma_j \right) = 0, \]  

(3.24)

since the double reaction of filaments of length \( i \) produces filaments of length \( 2i \). Similarly, for term \( E \), we have that

\[ iF_{ij} = \begin{cases} 
-2iC_j, & \text{when } i \text{ is a reactant in the mixed reaction } R_j, \\
2iC_j, & \text{when } i \text{ is a product in the mixed reaction } R_j,
\end{cases} \]

which gives:

\[ E = \left( \sum_{j \in J_{\text{mix}}} \sum_{i=1}^{\left\lfloor \frac{N}{2} \right\rfloor} \sum_{k=i+1}^{N-i} -2iC_j \varsigma_j - 2kC_j \varsigma_j + 2(i+k)C_j \varsigma_j \right) = 0. \]  

(3.25)

So inserting Equation (3.24) and (3.25) into Equation (3.23), we obtain
B = 0. Hence considering both \( A = 0 \) and \( B = 0 \), Equation (3.21) becomes

\[
\sum_{i=1}^{N} \frac{d\mu_i}{dt} = 0.
\]

Integrating with respect to \( t \) yields

\[
\sum_{i=1}^{N} i\mu_i(t) = k, \quad t \geq 0,
\]

where \( k \) can be uniquely determined by the initial condition \( (\mu_0, \sigma_0) \) as

\[
k = \sum_{i=1}^{N} i \left( \mu_0 + \frac{1}{2} \text{tr} \left( \frac{\partial^2 f_i(\mu)}{\partial \mu \partial \mu^T} \sigma_0 \right) \right) \in \mathbb{R}.
\]

Thus for the 2MA equations of the IF system, there is conservation of mass for the dynamics governing the means.

The mathematical results for the 2MA equations of IF assembly of the present study conclude with Proposition 3.8. It is worth noting that further investigations were carried out but proved to be inconclusive for the time being. From the result of Proposition 3.7, we first sought to prove the local stability of the identified equilibrium by means of a Jacobian linearisation analysis. However, by the nature of the reactions in the IF system, the Jacobian \( Df \) has entire zero-columns. Hence \( \text{rank}(Df) < N + N^2 \) and \( Df \) has zero eigenvalues, implying the equilibrium is non-hyperbolic. As such, no conclusion could be drawn.

We also focussed on proving the uniqueness of the equilibrium of Proposition 3.7. We conjecture that this equilibrium is unique, but our supposition
remains to be proven. If the equilibrium is indeed unique, we propose to study its stability by making use of the conservation equation of Proposition 3.8, as carried out in Proposition 3.5 for the RREs. However, it should be noted that the nonnegativity of solutions $\mu_i$ remains to be proven as it was for the RREs (Proposition 3.2) and, hence, modifications of the expression of the Lyapunov function may be necessary. We surmise that using the absolute value of the means would provide the stability result we seek.

Despite these limitations, the behaviours of the 2MA equations were assessable by means of numerical simulation. These results are highlighted in the upcoming section (particularly 3.4.1 and 3.4.2).

### 3.4 Comparisons

In this section, we treat the numerical results of the CME (SSA), the RREs, and the 2MA equations in three ways:

1. The models in the study of in vitro IF assembly are compared to the experimental results of [73] to investigate if they are in agreement with the data.

2. All three models are compared in order to juxtapose the similarities and the differences in each approach.

3. The SSA is compared to the RREs to contrast the stochastic framework to the deterministic.
For all numerical results of this section, the CME was simulated via the SSA, and both the RREs and the 2MA equations were simulated with the ODE solver Matlab ode15s [56]. Additionally, all simulations were run with the initial condition $(N_1(0), N_2(0), ..., N_N(0)) = (N, 0, ..., 0)$, where $N$ is the total number of populations. Throughout, the measures of comparison (relative error and sums of squared residuals or SSR) given in Section 3.1 are used. Experimental mean lengths and length distributions used in the present study were provided by the Divisions of Molecular Genetics and Biophysics of Macromolecules, German Cancer Research Center (DKFZ) Heidelberg, Germany.

### 3.4.1 Experimental Results

We begin with the comparison of experimental data to the CME. Having clarified the behaviour of the SSA in Section 3.1, we are now able to compare its results to experimental data from Portet et al., 2009 [73] to study the effects of stochasticity in the case of in vitro IF assembly, which lays at the core of the present study. As is routine in model to experimental comparisons, rate parameters used in [73] underwent extensive suitability testing to find best fits. Macroscopic rates $k_0d_{i,j}$ (referred to as $k_j$ in this study) were calculated using two parameters $k_0$, the intrinsic bimolecular rate constant, and $d_{i,j}$, a length-dependent proportionality factor. Underlying the specific values of each of the parameters $d_{i,j}$ are the hypotheses that the assembly process is diffusion-driven and filaments are rigid, rod-like objects. From these suppositions, parameter fitting was only undertaken for the intrinsic bimolecular rate constant $k_0$. 
We asked how we could use the results of [73] to study the system’s random behaviour. In fact, parameter values can be suitably scaled from the *macroscopic* level of mass action kinetic models while preserving the shapes of the macroscopic distributions [34]. To that end, Laurenzi, 2000 [51] shows that this scaling can be done according to the specific order of each of the reactions under consideration. Originally, based on the simple observation that the macroscopic rates are given in concentration per time, Gillespie, 1977 [21] proposed simply scaling by the experimental volume $\Omega$, so $C_j = \Omega k_j$. Later authors noted (see, for example, [24, 26, 52]), however, that it is common practice to express the deterministic rates per molar concentration per time, or $\text{moles} \cdot \text{(vol} \cdot \text{time})^{-1}$, so the microscopic rates should be given by $C_j = (k_j \cdot \text{vol}) \cdot N_A^{-1}$, with $N_A = 6.02 \cdot 10^{23}$ known as Avogadro’s number. To insist upon the necessity of using the parameters from deterministically fitted values, SSA runs using random rates of the same orders as the scaled reaction rates ($C_j = \text{rand} \cdot 10^{-5}$) were compared to the experimental data. This approach served to underline the fact that the deterministic fitting undertaken by Portet et al., 2009 [73] was crucial to conserving the experimental distribution shapes.

We therefore began our investigation of the stochastic influence on IF assembly using the fitted rate constants of [73] and scaled per the literature. We observed, however, that the reaction speeds were not of the appropriate time-scale for comparison to the experimental work. Interestingly, for the SSA, it was determined that rates in the order of $10^{-4}$ to $10^{-5}$ produced
dynamics with a time-scale similar to the experiment. This implied that the
deterministic rate constants $k_j$ of [73] needed to be multiplied by a factor
of $10^{-10}$. It is important to note, however, that no parameter fitting was
undertaken in the microscopic cases of consideration in the present research.

Knowing that the time-scale of the dynamics can be modified by multiplying
the reaction rates, we next varied the $C_j$ within the orders described above in
three multiples ($C_j = k_j \cdot 10^{-10}$, $C_j = 1.5 \cdot k_j \cdot 10^{-10}$, and $C_j = 2 \cdot k_j \cdot 10^{-10}$) to find the best agreement to the data amongst the three generated
dynamics. Parameter scaling results are given in Table 3.6, where SSA mean
lengths (MLs; number of ULFs) are compared to experimental results over
six time points (Figure 3.11). The most appropriate rates were determined
by comparing the mean relative errors of the MLs over the six time points
of interest. Standard deviation (SD) was also taken into account.
the deterministically scaled rates are illustrated in Figure 3.1. 1.5 - $k_j \cdot 10^{-10}$ were chosen to be of best fit over all times as highlighted in red. ML and relative error results for mean error. Relative errors are calculated using the experimental ML as a reference. Rates given by $C_j = 1.5 \cdot k_j \cdot 10^{-10}$ were chosen to be of best fit over all times as highlighted in red. ML and relative error results for mean error. Relative errors are calculated using the experimental ML as a reference.

### Table 3.6: Comparison of SSA and Experimental Data

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>SSA Mean Lengths (in number of ULFs)</th>
<th>Experimental Mean Lengths (in number of ULFs)</th>
<th>Relative Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.16</td>
<td>1.4631</td>
<td>1.3093</td>
<td>1.338</td>
</tr>
<tr>
<td>1.17</td>
<td>1.4842</td>
<td>1.3632</td>
<td>1.383</td>
</tr>
<tr>
<td>1.18</td>
<td>1.5053</td>
<td>1.3984</td>
<td>1.406</td>
</tr>
<tr>
<td>1.19</td>
<td>1.5265</td>
<td>1.4095</td>
<td>1.419</td>
</tr>
<tr>
<td>1.20</td>
<td>1.5477</td>
<td>1.4195</td>
<td>1.425</td>
</tr>
</tbody>
</table>

The table above compares the mean lengths from the SSA model to the experimental results for different times. The relative errors are calculated using the experimental mean lengths as a reference. The rates chosen for the SSA model are highlighted in red to indicate the best fit over all times.
CHAPTER 3. RESULTS

107

(a) SSA mean lengths using different rate constants

(b) Relative error of mean length results for SSA

Figure 3.11: **Comparison of SSA and Experimental Data**—Mean lengths from the SSA using different reaction rates are compared to the experimental means. Over long and short times, \( C_j = 1.5 \cdot k_j \cdot 10^{-10} \) were found to give the most appropriate time-scale. The experimental initial condition was \( N = 892 \) ULFs, whereas the CME was simulated for \( N = 446 \). Since we have shown that distributions from a \( k \)-fold decrease in the IC are statistically equal to those resulting from a \( k \)-fold increase in time (Section 3.1.2), simulation time was doubled for all models. The figures above reflect the data given in Table 3.6.
From the results of Table 3.6, we can see that over the six time points, rates given by $C_j = 1.5 \cdot k_j \cdot 10^{-10}$ provide the smallest mean relative error, coupled with a small standard deviation value. They were, accordingly, chosen as the rate constants best suited to the experimental ML values and were subsequently used in our latter investigations.

We next compared MLs and distributions obtained from the SSA, the RREs, and the 2MA equations to the experimental results (Table 3.7). First, for the CME/SSA, simulations at $t = 60s, 300s, 600s$, and $1200s$ over 5000 simulations were generated. Differences in final average distribution outputs of the SSA and the experimental results are shown in Figure 3.12.

We then compared the RREs to the experimental data of [73]. These results were calculated using the same time points and parameter rates ($C_j = 1.5 \cdot k_j \cdot 10^{-10}$) so we could contrast the stochastic model of the CME with one that is completely deterministic (RREs). This approach allows us not only to identify if stochasticity plays a role in the in vitro IF assembly system, but also, if it does indeed exist, what role it plays and when it is most important. These results are reflected in Figure 3.13.

Third, we compared distributions from the simulation of the 2MA equations to the experimental results. Owing to the system size of the 2MA equations ($N + N^2$ differential equations), simulations were carried out for $N = 50$. Unfortunately, this meant that a numerical side effect was produced when $t$ became too large, so distributions were unattainable for times
exceeding \( t = 600\)s. Nevertheless, three time points were compared to the results of [73], these being \( t = 60\)s, 300s, 600s. Distributions in proportions were compared to the experimental distributions and these results are shown in Figure 3.14.

(a) Mean Length Comparisons

<table>
<thead>
<tr>
<th>t (s)</th>
<th>Exp ML</th>
<th>SSA ML</th>
<th>( Err_{rel} ) (%)</th>
<th>RRE ML</th>
<th>( Err_{rel} ) (%)</th>
<th>2MA ML</th>
<th>( Err_{rel} ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>2.12</td>
<td>2.5941</td>
<td>22.36</td>
<td>2.2038</td>
<td>3.9517</td>
<td>2.1340</td>
<td>0.66</td>
</tr>
<tr>
<td>300</td>
<td>5.27</td>
<td>5.0636</td>
<td>3.92</td>
<td>5.2437</td>
<td>0.4997</td>
<td>5.0659</td>
<td>3.87</td>
</tr>
<tr>
<td>600</td>
<td>6.92</td>
<td>6.8325</td>
<td>1.26</td>
<td>6.8416</td>
<td>1.1323</td>
<td>6.4309</td>
<td>3.71</td>
</tr>
<tr>
<td>1200</td>
<td>10.93</td>
<td>9.6025</td>
<td>12.15</td>
<td>10.4990</td>
<td>3.9435</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

(b) Distribution Comparison

<table>
<thead>
<tr>
<th>t (s)</th>
<th>SSA</th>
<th>RRE</th>
<th>2MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>0.0318</td>
<td>0.0240</td>
<td>0.02641</td>
</tr>
<tr>
<td>300</td>
<td>0.0081</td>
<td>0.0090</td>
<td>0.0227</td>
</tr>
<tr>
<td>600</td>
<td>0.0206</td>
<td>0.0048</td>
<td>0.0154</td>
</tr>
<tr>
<td>1200</td>
<td>0.0046</td>
<td>0.0016</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 3.7: **Comparisons of Models to Experimental Data**-Using the adjusted reaction rates \( C_j = 1.5 \cdot k_j \cdot 10^{-10} \), results of the SSA, the reaction rate equations (RREs) and the two-moment approximation (2MA) equations are compared to the experimental results of [73]. The comparison of mean lengths (MLs) in (a) reports relative errors and shows all models relatively accurately reflect the experimental data. (b): From the small SSR results, each model is shown to differ little from the experimental data. In each case, the SSR decreases with an increase in time.
CHAPTER 3. RESULTS

Figure 3.12: Comparison of SSA and Experimental Data Using Rates $C_j = 1.5 \cdot k_j \cdot 10^{-10}$. On the left: differences in experimental and final average distributions of SSA (in proportions) over various time points. Differences are reduced with an increase in time as more filament lengths are produced (first column of Table 3.7(b)). On the right: experimental distributions and final average distributions of SSA over four time points. Each SSA average length distribution compares well with the experimental results of [73].
Figure 3.13: Comparison of RREs and Experimental Data Using Rates $C_j = 1.5 \cdot k_j \cdot 10^{-10}$. On the left: differences in experimental distributions and distributions of the RREs (in proportions) over various time points. Differences are reduced with an increase in time as more filament lengths are produced (second column of Table 3.7(b)). On the right: experimental distributions and RRE distributions over four time points. RRE distributions over all time points mimic the experimental results of [73].
Figure 3.14: **Comparison of 2MA and Experimental Data Using Rates** $C_j = 1.5 \cdot k_j \cdot 10^{-10}$ - On the left: differences in experimental distributions and distributions of the 2MA equations (in proportions) over various time points. Differences are reduced with an increase in time as more filament lengths are produced (third column of Table 3.7(b)). On the right: experimental distributions and 2MA distributions over three time points. Distributions from the 2MA equations demonstrate effects from the variability in the system but nonetheless represent the experimental data of [73] in a consistent manner over all three time points.
As a final illustration of the three models in comparison to the experimental results, Figure 3.15 shows the average length distribution of the SSA against the distributions from the simulation of the RREs, the 2MA equations and the data of [73] over short time \((t = 60s)\) and long time \((t = 600s)\).

All things considered, despite the lack of data fitting for the microscopic rate parameters, the SSA, the RREs, and the 2MA equations reflect the experimental data quite accurately. All three models provided reliable representations of the experimental data presented in [73] (Table 3.7). Since we wished to examine the validity of a stochastic approach to the modelling of IF assembly, and as the RREs generally provided the best representation (in relative error and in SSR), we can conclude that random behaviours are not of a significant importance to this specific system. In point of fact, this judgement could arise because of the peculiarities of the IF system, most notably the comparability of the reaction rate parameters. As shown in Figure 3.8(a), each reaction is almost equally likely. As Section 3.4.3 will show, when reaction probability rates exhibit large differences from one reaction to the next, the RREs behave differently than the CME and the 2MA equations. Furthermore, the parameters in [73] were fitted for the coagulation system analogous to the RREs, so undertaking parameter fitting specific to both the CME and the 2MA equations could modify our finding the RREs to be of best fit to the experimental data. As such, an important factor in future work will be the determination and fitting of rate parameters to ensure that reaction probabilities used in the microscopic case are actually reflective of the system. If fitted reaction rate probabilities exhibit large
discrepancies from one reaction to the next, noise can be deemed to be an important factor in the modelling of IF polymerisation or cast doubt onto the underlying hypothesis of the length dependency of the deterministic rates assumed in [73].
CHAPTER 3. RESULTS

Figure 3.15: **Comparison of all Models and Experimental Data**: Distributions from the simulation of all three models are compared with the experimental data over a short and long time points. Overall, all three models accurately represent the results of [73].
3.4.2 Modelling Approaches

As a second means of comparison, we examined the differences and similarities between the three models. While we expect all the models to compare favourably, the stochastic nature of the SSA and the 2MA equations should produce some variability in the results when compared to the RREs. For this investigation, rates $C_j = 1.5 \cdot k_j \cdot 10^{-10}$ scaled from the macroscopic system were used (Section 3.4.1). Mean lengths were compared over six time points and SSR was carried out to study the differences between the distributions (Table 3.8 and Figure 3.16). Since stochastic models are known to capture transient dynamics better than their deterministic counterparts [86], the study of this section concentrates on short-time points (from 5s to 600s) as they are the most demonstrative of these transitory kinetics.
(a) Mean Lengths

<table>
<thead>
<tr>
<th>t (s)</th>
<th>SSA</th>
<th>RRE</th>
<th>2MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.2121</td>
<td>1.0963</td>
<td>1.0914</td>
</tr>
<tr>
<td>10</td>
<td>1.3978</td>
<td>1.1865</td>
<td>1.6024</td>
</tr>
<tr>
<td>30</td>
<td>1.9733</td>
<td>1.6577</td>
<td>1.6021</td>
</tr>
<tr>
<td>60</td>
<td>2.5941</td>
<td>2.2038</td>
<td>2.5864</td>
</tr>
<tr>
<td>300</td>
<td>5.0636</td>
<td>5.2437</td>
<td>5.0659</td>
</tr>
<tr>
<td>600</td>
<td>6.8325</td>
<td>6.8416</td>
<td>6.6636</td>
</tr>
</tbody>
</table>

(b) Distribution Comparison

<table>
<thead>
<tr>
<th>t (s)</th>
<th>CME vs RRE</th>
<th>CME vs 2MA</th>
<th>RRE vs 2MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.354·10^{-4}</td>
<td>1.71·10^{-4}</td>
<td>6.32·10^{-4}</td>
</tr>
<tr>
<td>10</td>
<td>2.32·10^{-4}</td>
<td>3.25·10^{-4}</td>
<td>7.93·10^{-4}</td>
</tr>
<tr>
<td>30</td>
<td>3.2·10^{-4}</td>
<td>6.01·10^{-4}</td>
<td>1.39·10^{-4}</td>
</tr>
<tr>
<td>60</td>
<td>9.4·10^{-4}</td>
<td>9·10^{-4}</td>
<td>4.2·10^{-4}</td>
</tr>
<tr>
<td>300</td>
<td>2.36·10^{-4}</td>
<td>9·10^{-4}</td>
<td>1.05·10^{-4}</td>
</tr>
<tr>
<td>600</td>
<td>9·10^{-4}</td>
<td>8.9·10^{-4}</td>
<td>4·10^{-4}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>μ(SSR)</th>
<th>σ(SSR)</th>
<th>μ(SSR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0160</td>
<td>0.0045</td>
<td>0.0147</td>
</tr>
</tbody>
</table>

Table 3.8: Model Comparison-In (a), mean lengths (in number of ULFs) are shown to compare favourably across all time points. (b) shows the pairwise comparison of distributions (see Figure 3.16). Results in red highlight the model demonstrating the smallest SSR over a given time point: results from the 2MA compare well to those of the two other models. Overall, from the mean of the SSR (μ(SSR)), the CME and the 2MA equations are shown to be closest, a natural conclusion as the 2MA derive from the CME.
Figure 3.16: Model Comparison—Distributions in proportions of the three models are compared over six time points. The 2MA equations were simulated for \( N = 50 \) whereas the SSA and the RREs were simulated for \( N = 446 \).
CHAPTER 3. RESULTS

The results of this section demonstrate that all three models compare similarly to one another in mean length and in distribution (Table 3.8; Figure 3.16). That being said, from the middle column of Table 3.8(b), it is not surprising that the CME and the 2MA equations compared most favourably as they both account for the system’s noise and furthermore, the 2MA equations are a second moment truncation of the probability distributions of the CME. It is worth noting that over short times in particular, the distributions of all three models exhibit greater differences, thereby demonstrating the need to account for transient dynamics. As mentioned, stochastic models are more readily applicable to this study of short-term behaviour [26]. Hence, stochastic modelling is a valuable approach in the study of short-term behaviours of reaction and assembly systems.

3.4.3 Frameworks

Finally, as one of the interests of the present study is to explore how the presence of noise can influence assembly models, we are interested in comparing the numerical solutions of the RRE to the average length distributions from SSA runs in an effort to characterise stochastic effects. As described in Section 2.3, the nonlinear nature of most assembly systems implies that the RREs can only approximate the mean behaviours in the system. To study how these approximations can potentially produce deviations, time points in this section were chosen to reflect a mix of short and long times. For both models, the total number of populations was given by \( N \) and rate parameters were given by \( C_j = 1.5 \cdot k_j \cdot 10^{-10} \). The SSA was simulated over 5000 simulations runs. Mean lengths of each of the models are given in Tables
3.7(a) and 3.8(a). From the SSA we can obtain not only mean population numbers (as in the RRE) but the per population variance (Figures 3.17 and 3.18). In addition, to examine the behaviour of both models using wholly random rates, $C_j = \text{rand} \cdot 10^{-4}$ were used in each case, and the results are given in Figure 3.19.
CHAPTER 3. RESULTS

(121)

(a) \( t = 5 \)  
(b) \( t = 10 \)  
(c) \( t = 30 \)  
(d) \( t = 60 \)

Figure 3.17: Comparison of SSA to RRE - Average length distributions with standard deviations per filament length from the simulation of the SSA are compared to distributions from the simulation of the RRE over short times. The differences between the stochastic model SSA and the deterministic RREs are more pronounced over shorter times, and attenuate as time increases.
Figure 3.18: Comparison of SSA to RRE. Average length distributions with standard deviations per filament length from the simulation of the SSA are compared to distributions from simulation of the RRE. The x-axis represents filament lengths in number of ULFs while the y-axis reports the proportion of filaments. Increases in time induce smaller standard deviations as more filament lengths are incorporated and created in the system. Over longer times, the differences between the average length distributions of the SSA and distributions from the RRE are reduced.
Figure 3.19: **Comparison of SSA to RRE**-Average length distributions and standard deviations from the simulation of the SSA are compared in proportions to distributions from the simulation of the RRE using random rates $C_j = \text{rand} \cdot 10^{-4}$. Large differences in distributions are observed over both short and long times, which was not observed using rates scaled from the model of [73], as seen in Figures 3.17 and 3.18.
Figures 3.17 and 3.18 show that for the IF system, stochasticity plays a small role, as distributions from the SSA generally compare so favourably to those of the completely deterministic RREs. There exists, nonetheless, differences in distributions and mean lengths, especially over short time. However, long term behaviours of the SSA agree with those of the RREs (Figure 3.18). As previously discussed, the regularities in the rate parameters given by $C_j = 1.5 \cdot k_j \cdot 10^{-10}$ diminish the role noise can play in the IF system. However, the effects of stochasticity are particularly highlighted when studying probability rate constants exhibiting large differences from one reaction to the next (Figure 3.19) and the RREs were shown to differ significantly from the SSA in this case.

The results of Section 3.4 allow us to conclude on the modelling approach undertaken in this thesis as follows:

1. The SSA and the 2MA equations were shown to agree with the experimental results of in vitro assembly of IFs [73] over all times considered, implying that stochasticity can be factored into models of IF assembly. However, the RREs are nonetheless able to accurately capture the system’s dynamics.

2. The models were shown to accurately represent each other’s results. The SSA and the 2MA equations were found to agree more consistently on average. This conclusion is natural because the 2MA equations are derived directly from the CME. However, it was nonetheless important to investigate their agreement because the 2MA equations are approx-
imations of the first two moments of the probability distributions of the CME. Further, both provide results on the variability between populations, something that is inaccessible in the case of the RREs.

3. Over longer times, differences in SSA and RREs distributions were found to decrease when the system is considered with relatively comparable reaction rates. Conversely, the results of SSA and the RREs demonstrate significant differences in an entirely stochastic system when reaction rates exhibit large differences from one reaction to another.

From the results detailed above, we have illustrated the importance of the study of stochasticity in assembly models, with a particular interest to the IF system. In the following chapter, we discuss the major findings of the present research.
Chapter 4

Discussion and Remarks

In the preceding chapters, we have outlined mathematical methods to include the random behaviours present in reaction systems and detailed their application to the assembly system of intermediate (IFs). In what follows, we now turn to a discussion of the research approaches of this thesis. This chapter is divided as follows:

1. Treatment of the mathematical foundations and results of the three models under consideration.

2. An examination of the computational approach of the present study and its numerical results.

3. Notes on the comparison of the models to the experimental data.

From these topics, we will then provide some global concluding remarks on the modelling approach studied in this thesis. The main results of each of the models are recapitulated in Figure 4.1.
CHAPTER 4. DISCUSSION AND REMARKS

Numerical results of all models agree with each other

Figure 4.1: The major results of the simulation and analysis of the chemical master equation (CME), the reaction rate equations (RREs), and the two-moment approximation (2MA) equations.
4.1 Treatment of Mathematical Results

Mathematical modelling provides significant advantages in the study of cellular phenomena as it can not only serve to confirm experimental results, but can also explain counterintuitive results and discover previously undiscovered relationships and behaviours [2]. While models are necessarily simplifications of the true processes in question, they nonetheless offer a depth of understanding that may be unobtainable by experimental methods alone [65]. As such, mathematical modelling is of crucial importance to future investigations of the cell and its organelles and processes. Of particular significance going forward is the identification and explanation of the variability so often observed within biochemical reaction systems [82]. To that end, the present study presented two methods suited to the investigation of stochasticity within assembly systems, and contrasted these models with a purely deterministic approach. We began by setting the foundations of the chemical master equation (CME), which studies the evolution of probability sets. If solved, its solutions explain the full probability distribution of all states at all times [82]. However, as a direct result of this complete description, the dimensions of the CME and its solutions quickly become intractable making mathematical analysis impossible except in a few specific cases [42].

In an effort, then, to mathematically study the noise inherent to biological systems, we turned to the two-moment approximation (2MA) equations which truncate the full probability distribution of the CME to the first two moments in the system [84]. Since the 2MA equations consider only sec-
ond order fluctuations between population states, the ordinary differential equations (ODEs) describing their time-evolutions become much more manageable. From the 2MA equations of intermediate filament (IF) assembly, we have proven the existence and uniqueness of solutions, identified one equilibrium point, and shown there is conservation of mass for the equations governing the means. Further, the 2MA equations were shown to be analogous to an established result from the theory of statistical thermodynamics known as the Fluctuation-Dissipation Theorem [69]. This confirms the deeply rooted principles of the 2MA equations in various disciplines ultimately underscoring their potential impact in the study of assembly events.

It would be remiss to fail to point out that the classic mass-action approach to reaction and polymerisation systems has quantitatively been shown to be a successful modelling framework [65]. As such, the use of reaction rate equations (RREs), though they fail to account for the dynamics of the system’s noise, is validated under certain hypotheses. And since the RREs discount the intrinsic variability between populations, they are defined on even smaller dimensions than the 2MA equations, which makes the use of the powerful theory on ODEs even more accessible. For the RREs we have shown the existence and positivity of solutions, the conservation of mass within the system, and the global asymptotic stability of a unique equilibrium point via a Lyapunov function. Of particular interest is that this unique equilibrium point is analogous to the one found in the case of the 2MA equations and observed in the long-time simulation of the SSA, which in all three cases is coherent with the experimental result of the system’s aggregation.
into one long filament when given sufficient time to react. However, the study of the system’s transient behaviours is generally more readily accessible by use of the SSA and the 2MA equations, by virtue of their stochastic nature [26] and furthermore, information on the system’s variability is not available from the RREs but can be obtained from both the SSA and the 2MA equations.

Throughout the present study, we have highlighted the importance of the cautious selection of any potential model in the study of assembly. Mathematical models in cellular biology offer many advantages; often the mere process of modelling is a crucial step to revealing a more intimate understanding of the question at hand [65]. Thus the incorporation of stochasticity into reaction models at the outset of its construction may already shed light on the involved processes, and remains an important factor of consideration when examining cellular structures.

### 4.2 Computational Approaches

As we have seen, the characterisation of stochastic dynamics with the CME can provide a great deal of insight into the behaviour of reaction systems and their inherent statistical noise [22, 59, 82, 86]. Though solutions to the CME are rarely obtainable [13, 42], Gillespie introduced the stochastic simulation algorithm, or SSA as a means to simulate the CME using Monte-Carlo methods [21]. The exact nature of the SSA and its accurate representation of the most probable configuration of the system described by the CME at
a given time explains its prevalence in the study of random behaviours in cellular processes [36,57,74,82,86].

The reasons the SSA exactly reproduces the most probable configuration of the system described by the CME at a given time were outlined in Section 2.2.2, and can be summarised by noting that the properties of the two random variables (RVs) \( \tau \) (the time to the next reaction) and \( j \) (the next reaction’s index) are foundational to the technique since we repeatedly sample from their probability density functions [67]. And since the SSA is not an approximation of solutions, as is the case in the simulation of ODEs by techniques like the commonly used Runge-Kutta method, the CME remains a noteworthy modelling approach. Applying the SSA for the simulation of the CME, we were able to compare all three models numerically. An outline of the results from this computational analysis follow.

Before comparing the results of all three models, it was necessary to explore the intricacies of the SSA (Section 3.1). Accordingly, we sought to answer three questions related to the use of the SSA for the simulation of the CME:

1. How many simulation runs of the SSA produce a tolerable convergence and allow the calculation of a reliable average length distribution?

2. How do changes to the total number of populations affect the SSA results?

3. How do changes to parameters affect the numerical solutions of the
SSA?

The results of Sections 3.1.1, 3.1.2, and 3.1.3 provide the following answers:

1. More simulations are more desirable to ensure an acceptable level of convergence.

2. Changes to the total number of populations affect average length distributions, but proportional increases to the total number of ULFs coupled with proportional decreases in time produce statistically equivalent average length distributions.

3. Modifying reaction rates profiles can change the shape of distributions. Further, rate scaling enables us to adjust the time-scale of the dynamics.

Having identified these behaviours, we then turned to the comparison of the three models to contrast their numerical results. Further, we analysed the differences the addition of stochasticity had on assembly models. The results of Section 3.4.2 showed that all three models were comparable in mean lengths and in distributions through either relative error or sums of squared residuals (SSR) measures. Taking into account the average of the SSR, we found that distributions from the 2MA equations and the CME were very close to each other, which was somewhat expected considering that the 2MA equations derive from the moment closure of the CME. We also observed that all models differed most over shorter times. This speaks to the difficulty in capturing short-time transient behaviours in assembly systems, which can be more accurately depicted using stochastic techniques [26, 86]. Moreover,
both the SSA and the 2MA equations allow us to access information on the variability in the system, which is unavailable when applying the RREs.

Section 3.4.3 demonstrated key points in the investigation of differences between a stochastic approach and one that is completely deterministic. Using the rate parameters scaled from [73], the SSA and the RREs, though they exhibited small dissimilarities, were generally found to compare quite well to one another. However, when using randomly derived reaction rate probabilities, the results of the SSA and of the RREs differed greatly. This indicates the CME is more suitable in systems exhibiting a mix of rare reactions and highly probable reactions, as is common in cellular behaviour [49], which again underlines the importance of appropriate model selection [88]. Moreover, the potential for observing regularities or peculiarities of a given system, like the aforementioned rare reactions, during model construction demonstrates the observation that conclusions can emerge both during the construction phase and the analysis stage of modelling in the biological sciences [65].

4.3 Experimental Comparison

Ultimately, a primary motivation of the present research was the application of stochastic models to the experimental results of Portet et al., 2009 [73] to investigate the role of stochasticity in IF polymerisation. Before conducting this numerical investigation, it was crucial to note that no data fitting was undertaken in the microscopic case of interest in this study, as was car-
ried out in [73]. This approach is not invalid as several authors underline the soundness of parameter scaling from deterministic models [21, 24, 26, 51]. However, upon examination of the macroscopic rates scaled according to the literature, we observed the necessity of further adjustments in the reaction rates to align the numerical and experimental time-scales. These modifications were obtainable through the results of Section 3.1.3 which showed that multiplying reaction rates by a given factor fine-tunes the time-scale of the dynamics.

Having considered the appropriate parameters, we then turned to comparing the mean lengths and distributions of each model to the experimentally derived mean length and distribution results. Overall, all three models mimicked the data. It was shown that the 2MA equations compare favourably when the CME did as well (and vice versa). This is to be expected as the 2MA can be viewed as a smaller-size representation of the CME’s probability distributions. However using sums of squared residuals, we showed the RREs to best fit to the data. This again underscores the importance of the regular reaction rate probabilities used in our investigations. Since each reaction is generally as probable as any other, the system does not exhibit drastic parameter effects, which stresses the necessity of parameter fitting in the case of microscopic models to verify the findings of this study. Furthermore, the data fitting undertaken in [73] was specific to the coagulation equation description analogous to the RREs. As such, we can surmise that the scaled reaction rates contribute to the result of a good fit to the data in the case of the RREs. Nonetheless, finding the RREs to be
most reflective of the experimental data does not necessarily invalidate the use of stochastic models for IF assembly. Given that models accounting for random behaviours can more readily capture transient dynamics, the use of the CME and the 2MA equations is a reasonable approach to the modelling of short-time responses.
Chapter 5

Conclusion

The present research investigated the inclusion of random behaviour in classes of reaction systems and then applied these approaches to the assembly of intermediate filaments (IFs). Using techniques from the theory of dynamical systems, probability theory, matrix theory, and analysis, we have demonstrated how sets of ODEs can be used to study the inherently erratic behaviour observed in biological systems. The CME, a tool frequently relied upon for the study of reaction systems and assembly models, is not only a valid modelling procedure, but also forms the basis of the 2MA equations, while the RREs can be viewed as the completely deterministic version of these two models.

In the preceding material, we have outlined the derivations of the these modelling approaches and have mathematically and/or numerically investigated their behaviour. This, in turn, allowed for the comparison of all three to each other and for the application of all three to the IF system of inter-
As mentioned, the assembly of IFs is an important process to study for many reasons. Most notably, as IFs are vital to many cellular processes, they have a direct effect on several devastating diseases [68]. Since cells demonstrate random processes, the inclusion of stochasticity in mathematical models describing their behaviour is essential to capture a full picture of cellular comportment [82], especially when considering transient responses. As shown in Section 3.4.1 and discussed in Section 4.3, all models compare very well to the experimental data, which was best represented by the RREs. Two factors can account for this finding. First, the large system and the frequent and almost equal probability of all reactions in the IF system under the hypotheses of [73] contribute a certain regularity. Second, deterministic rates used here were obtained by data fitting undertaken for the analogous coagulation equations in [73] meaning the good fit of the RREs may be attributed to this procedure. Randomness, however, remains an important consideration in assembly models, especially when taking into account possible rare reactions. It is crucial to note that, as with any modelling technique, the application of the CME, RREs, and the 2MA equations to IF assembly necessitate simplifications. The assembly considered in this thesis occurs in vitro only. Quite obviously, the in vivo assembly of IFs is drastically different [71]. In cells, IFs form complex network structures and can disassemble. The models considered herein discount both of those occurrences.

In terms of IFs, it will be important to mathematically study the formation of their networks, further to the work of [72]. It remains possible to apply the stochastic techniques emphasised in this thesis to such problems,
as has been done for reaction-diffusion processes (which are important for the initiation of IF networks in vivo) [14], for tumour proliferation and flow through vascular networks [58], and for the simulation of polymer networks [64]. Work on IF networks will help to complete the modelling of the cytoskeleton as a whole and, therefore, provide a detailed and mathematically supported global view of its structure and function within the cell. This in turn increases our understanding of IFs and the cytoskeleton, which only strengthens efforts in disease prevention and management, for example.

To conclude, we found the incorporation of stochasticity into problems in cellular biology to be an important aspect of consideration. There has been an increasing interest in this modelling approach in the literature, beginning in the 1960s with applications to chemical reaction systems [59] to the present day, in studies of microtubules, one of the IF’s counterparts [57]. In the coming years, as experimental work in cell biology is increasingly complemented by mathematical models, it will become progressively more important to account for the random behaviour we observe in the smallest of organelles to the cell as a whole. This will maximise the reliability of results and therefore underscores the significance of including stochasticity in models of cellular processes. Ultimately, mathematical modelling in cell biology is an extremely powerful tool of investigation and one that continues to evolve as its use becomes increasingly prevalent.
Relevant Definitions and Theorems

The present study drew from definitions and theorems from various disciplines within mathematics. These theoretical tools are presented herein and are organised by area of study.

5.1 Vector Calculus and Matrix Theory

We begin with definitions pertaining to the differentiability of vector valued functions. As is our convention, we refer to a $1 \times N$ line vector as an $N$-vector, and specify dimensions for column vectors.

**Definition 5.1.** [6] Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and suppose every component $f_i$ $(i = 1, \ldots, m)$ of $f$ admits all first partial derivatives $\frac{\partial f_i}{\partial x_j}$, $j = 1, \ldots, n$ so that to have the vector

$$\nabla f_i = \left( \frac{\partial f_i}{\partial x_j} \right)_{1 \leq j \leq n} = \left( \frac{\partial f_i}{\partial x_1}, \ldots, \frac{\partial f_i}{\partial x_n} \right),$$

here written as a row vector. The $m \times n$ matrix
Df = \left( \frac{\partial f_i}{\partial x_j} \right)_{1 \leq i \leq m, 1 \leq j \leq n} = \begin{pmatrix} \nabla f_1 \\ . \\ . \\ . \\ \nabla f_m \end{pmatrix}

is called the Jacobian matrix of f. The notation Df(a) indicates the Jacobian matrix is evaluated at the point a ∈ \mathbb{R}^n.

**Definition 5.2.** [18] Suppose f : \mathbb{R}^n \rightarrow \mathbb{R} is a real-valued function of class \(C^2\) on some open set S ⊂ \mathbb{R}^n. The \(n \times n\) matrix \(H\), called the Hessian of \(f\) is:

\[ H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \ldots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \ldots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}, \]

where \(\frac{\partial_i f}{\partial x_i}\) is the partial derivative of \(f\) with respect to \(x_i, i = 1, \ldots, n\). An equivalent vectorial definition of the Hessian is given by:

\[ D^2 f = H = \frac{\partial^2 f}{\partial x \partial x^T}, \]

where \(x = (x_1, \ldots, x_n)\). The notation \(D^2 f(a)\) denotes the evaluation of \(D^2 f\) at the point \(a \in \mathbb{R}^n\).
Equipped with the two definitions above, we can now define the Taylor series of a function, particularly useful in the derivation of the two-moment approximation equations.

**Definition 5.3.** [18, 50] For any infinitely differentiable function \( f(x) : \mathbb{R} \rightarrow \mathbb{R} \), the Taylor series centred at \( x = x_0 \in \mathbb{R} \) is given by

\[
\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n,
\]

where \( f^{(n)}(x_0) \) is the \( n \)th derivative of \( f \) evaluated at \( x_0 \). An \( n \)th order finite truncation of the Taylor series of \( f(x) \) about \( x_0 \) is given by:

\[
f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n(x)
\]

with remainder satisfying:

\[
|R_n(x)| \leq M_n \frac{(x-x_0)^{n+1}}{(n+1)!}
\]

for positive constants \( M_n \), which tends to zero with an increase in \( n \).

For \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), we can express the second order truncation of the Taylor series of \( f \) about \( a \in \mathbb{R}^n \) as [18]:

\[
f(x) = f(a) + \sum_{j=1}^{n} \partial_j f(a)(x_j - a_j) + \frac{1}{2} \sum_{i,j=1}^{n} \partial_i \partial_j f(a)(x_i - a_i)(x_j - a_j) + R_2(x),
\]
or, alternatively in matrix form as:

\[ f(x) = f(a) + Df(a)(x - a) + \frac{1}{2}(x - a)^T D^2 f(a)(x - a). \]

for column vectors \( x \) and \( a \), and \( Df \) as in Definition 5.1 and \( D^2 f \) as in Definition 5.2.

An important result related to the Taylor series is the notion of analytical functions, defined below. Theorem 5.1 then provides a property relating two such functions.

**Definition 5.4.**[50] If the series \( \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \) converges towards \( f(x) \) for all \( x \in (-\delta + x_0, \delta + x_0) \), \( \delta > 0 \), we say that \( f \) is analytic in \( (-\delta + x_0, \delta + x_0) \).

**Theorem 5.1.**[17] The sum and product of functions analytic on a common domain are again analytic.

Definitions 5.1 to 5.4, with Theorem 5.1 were frequently applied in the derivation of the 2MA equations. Of particular importance was the finite Taylor series given in Definition 5.3, as it allowed for a truncation of the moments of the chemical master equation (CME). Turning to matrix theory, of last consideration in this section is a definition for a product of given matrices which was of particular importance to the definition of propensity functions.

**Definition 5.5.** Let \( A, B \in \mathbb{R}^{m \times n} \). The Hadamard product \( A \odot B \in \mathbb{R}^{m \times n} \)
is the result of the element-wise multiplication of \( A \) and \( B \). Thus \((A \circ B)_{ij} = A_{ij} \cdot B_{ij}\).

## 5.2 Combinatorics and Number Theory

The two basic results of combinatorics were employed in this manuscript.

**Definition 5.6.** [85] A combination of \( r \) objects from a collection of \( m \) objects is any unordered arrangement of \( r \) distinct objects from the \( m \) objects. The number of possible combinations of \( r \) objects that can be formed from a collection of \( m \) objects is denoted \( \binom{m}{r} \) and is given by

\[
\binom{m}{r} = \frac{m!}{r!(m-r)!}
\]

for \( m \geq r, m, r \in \mathbb{N} \).

The above definition can be extended as follows,

**Definition 5.7.** For \( m \in \mathbb{Z} \) and \( r \in \mathbb{N} \), we define

\[
\binom{m}{r} = 0
\]

if \( m < r \).

Further, a frequently utilised result is given below.

**Definition 5.8.** [80] The floor function \( [x] \) is the greatest least integer function, satisfying the following property:
\[ \left\lfloor \frac{x}{2} \right\rfloor = \begin{cases} \frac{x}{2} : x \in 2\mathbb{N} \\ \frac{x-1}{2} : x \notin 2\mathbb{N} \end{cases}, \]

where \( 2\mathbb{N} \) is the set of even integers.

### 5.3 Probability Theory

One of the major focuses of the present study is the incorporation of stochasticity into assembly models. Stochasticity directly refers to random variables (RVs) and accordingly, its theory is rooted in probability theory. Some foundational principles from this discipline are discussed within this section.

**Definition 5.9.** [85] Let \( E \) be an event and let \( P(E) \) denote its probability. For \( n \) independent repetitions of the random experiment, let \( n(E) \) denote the number of times that event \( E \) occurs. Then the frequentist interpretation of probability is

\[ \frac{n(E)}{n} \approx P(E), \quad \text{for large } n. \]

From this probability \( P(E) \), we have the following definitions.

**Definition 5.10.** [85] For each event \( E \),

\[ P(E) = 1 - P(E^C), \]

where \( E^C = \{x|x \notin E\} \) is called the complement of \( E \). In words, the probability that an event occurs equals 1 minus the probability that it doesn't occur.
**Definition 5.11.** [85] Two events $A$ and $B$ are said to be independent events if

$$P(B \cap A) = P(A)P(B).$$

In words, two events are independent if their joint probability equals the product of their marginal probabilities. If $A$ and $B$ are independent events, knowing that event $A$ occurs doesn’t affect the probability of occurrence of event $B$.

Considering now probabilities as they relate to random variables, consider the definitions given below.

**Definition 5.12.** [85] The RV $X$ has a probability density function (PDF) $f_X(x)$ defined as

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx.$$

**Definition 5.13.** [85] Let $X$ be a RV. Then the cumulative distribution function of $X$, $F$, is the real-valued function defined on $\mathbb{R}$ by

$$F(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

Since the simulation method of the CME known as the stochastic simulation algorithm (SSA) utilises the generation of random numbers from specific PDFs, it is important to note the following specific distributions of a RV.

**Definition 5.14.** [85] A continuous RV $X$ is called a uniform RV if, for some finite interval $(a, b)$ of real numbers, its value is equally likely to lie
anywhere in that interval or, equivalently, its PDF is constant on that interval and 0 elsewhere. We say that $X$ has a uniform distribution on the interval $(a,b)$ or that $X$ is uniformly distribution on the interval $(a,b)$. For convenience, we sometimes write $X \sim U(a,b)$ to indicate that $X$ has the uniform distribution on the interval $(a,b)$.

**Definition 5.15.** [85] A continuous RV $X$ is called an exponential RV if it has a PDF of the form

$$P(X < x) = f_X(x) = \lambda e^{-\lambda x}, \quad x > 0$$

and $f_X(x) = 0$ otherwise, where $\lambda$ is some positive real number. We say that $X$ has an exponential distribution with parameter $\lambda$ or that $X$ is exponentially distribution with parameter $\lambda$. For convenience, we sometimes write $X \sim E(\lambda)$ to indicate that $X$ has an exponential distribution with parameter $\lambda$. $X$ has an expected value $E[X] = \frac{1}{\lambda}$ and variance $Var[X] = \frac{1}{\lambda^2}$.

As has been treated throughout, the 2MA equations are in fact a truncation to the first two moments of the probability distributions of the CME. Consider the expectation given by

**Definition 5.16.** [85] Let $X$ be a discrete RV with sample space $\omega$ (the set of all possible outcomes of $X$). The expected value of $X$, denoted $E[X]$, is defined by

$$E[X] = \sum_{x \in \omega} xP(X = x),$$

and is also referred to as the first central moment, or expectation. Let $X$
and $Y$ be discrete RVs defined on the same sample space and having finite expectation, and let $c$ be a real number. Then the following relations hold.


b) The RV $cX$ has finite expectation and $E[cX] = cE[X]$.

Further, for an arbitrary function $f(X)$ of $X$

c) $E[f(X)] = \sum_{x \in \omega} f(x)P(X = x)$.

In general, the moment is defined as follows.

**Definition 5.17.** [85] Let $X$ be a RV with finite expectation with sample space $\omega$. If $X - E[X]$ has a finite $r^{th}$ moment, we define the $r^{th}$ central moment of $X$ to be the $r^{th}$ moment of the RV $X - E[X]$:

$$E[(X - E[X])^r] = \sum_{x \in \omega} (x - E[X])^r P(X = x).$$

In the case of the first two moments of interest to the 2MA equations, we are particularly interested by the second moment.

**Definition 5.18.** [85] Let $X$ and $Y$ be RVs defined on the same sample space and having finite variances. The covariance of $X$ and $Y$, denoted $\text{Cov}(X,Y)$ is

$$\text{Cov}(X,Y) = E[(X - E[X])(Y - E[Y])] = \text{Cov}(Y,X),$$

and is also referred to as the second central moment. It is frequently calculated by $\text{Cov}(X,Y) = E[XY] - E[X]E[Y]$. The measure of covariance
between a RV $X$ and itself is known as the variance and is written as

$$\text{Cov}(X, X) = \text{Var}(X) = E[(X - E[X])^2],$$

where $\sqrt{\text{Var}(X)}$ is defined to be the standard deviation of $X$.

Further, consider the following result for the expectation of a quadratic (bilinear) form.

**Definition 5.19.** [66] Let $X = (X_1, X_2, ..., X_p)^T$ be a random vector and $A$ be a $p \times p$ symmetric matrix of constants. Let $E[X] = \mu$ and let $\text{Cov}(X) = \sigma$, where $\sigma$ is the covariance matrix with entries $\sigma_{ik} = \text{Cov}(X_i, X_k)$. Then

$$E[X^TAX] = \text{tr}(A\sigma) + \mu^TA\mu,$$

where $\text{tr}(\cdot)$ denotes the trace of a square matrix.

Note the following equivalent expression:

$$E[X^TAX] = \text{tr}(A\sigma) + \mu^TA\mu$$

$$\iff \text{tr}(A\sigma) = E[X^TAX] - \mu^TA\mu$$

$$= E[X^TAX - \mu^T\mu]$$

$$= E[(X - \mu)^T(AX - A\mu)]$$

$$= E[(X - \mu)^TA(X - \mu)].$$
Finally, the derivation of the 2MA equations from the CME drew upon the following result of Engblom, 2006 [13]. To begin, consider the two following definitions from functional analysis.

**Definition 5.20.** [41] The support of a function \( f(x) \) is the closure of the set \( \{ x \in X | f(x) \neq 0 \} \). It is denoted \( \text{supp}(f) \). If \( \text{supp}(f) \) is a compact subset of the domain \( X \), then \( f(x) \) is said to have compact support.

**Definition 5.21.** [41] The function \( f(x) \) is a test function if it is continuously infinitely differentiable (i.e. \( f(x) \in C^\infty \)) and it has compact support.

From the definition of a test function, Theorem 5.2 was used to derive the 2MA equations from the CME.

**Theorem 5.2.** [13] Let \( P \) be a probability distribution satisfying the chemical master equation (CME):

\[
\frac{dP(n,t|n_0,t_0)}{dt} = \sum_{j=1}^{M} [a_j(n - S_{-j})P(n - S_{-j},t|n_0,t_0) - a_j(n)P(n,t|n_0,t_0)]
\]

(for simplicity, let \( P(n,t|n_0,t_0) = P(n,t) \)). Then for a given test function \( T : Z^{N+} \rightarrow \mathbb{R} \),

\[
\sum_n T(n)\frac{dP(n,t)}{dt} = \sum_{j=1}^{M} E[(T(N + S_{-j}) - T(N))a_j(N)].
\]

### 5.4 Theory of Dynamical Systems

The theory of dynamical systems is rich with powerful results allowing us to study asymptotic behaviour of these systems, in addition to other important
properties. Since all of the models discussed in the present studies are ODEs, the following definitions and theorems were necessary for their mathematical analysis.

**Definition 5.22.** [27] Let $D$ be an open subset of $\mathbb{R}^n$. A dynamical system on $D$ is the triple $(D, \mathbb{R}, s)$, where $s : \mathbb{R} \times D \rightarrow D$ is such that the following axioms hold:

1. (Continuity): $s(\cdot, \cdot)$ is continuous on $D \times \mathbb{R}$ and for every $t \in \mathbb{R}$, $s(\cdot, x)$ is continuously differentiable on $D$.

2. (Consistency): $s(0, x_0) = x_0$ for all $x_0 \in D$.

3. (Group property): $s(\tau, s(t, x_0)) = s(t + \tau, x_0)$ for all $x_0 \in D$ and $t, \tau \in \mathbb{R}$.

From the above definition, we can define the flow of a dynamical system.

**Definition 5.23.** [27] The map $s(\cdot, \cdot)$ is called the flow of $(D, \mathbb{R}, s)$. For a given $s(t, x_0), t \geq 0$, we refer to $x_0 \in D$ as an initial condition of $(D, \mathbb{R}, s)$.

Given $t \in \mathbb{R}$, $s(t, \cdot) : D \rightarrow D$ is written $s_t(x)$ or $s_t$.

Of particular interest in this study are ordinary differential equation initial value problems (IVPs) as defined below.

**Definition 5.24.** [27] Let $x(t) \in D$ and $f : \mathbb{R} \times D \rightarrow \mathbb{R}^n$. The dynamical system

$$\frac{dx(t)}{dt} = f(t, x(t)), \quad t \in \mathbb{R},$$

(5.1)

subject to $x(t_0) = x_0$ is called an initial value problem.
By the following theorem, we are assured of the existence and uniqueness of solutions to IVPs.

**Theorem 5.3.** [1] Assume that \( f : \mathbb{R} \times D \rightarrow \mathbb{R}^n \) and \( \frac{\partial f}{\partial x_i} \) (for \( i = 1, ..., n \)) are continuous functions in \( \mathbb{R}^n \). Then a unique solution exists to Equation (5.1) for any initial value \( x_0 \in D \).

In the interest of finding positive solutions to IVPs, we define positively invariant domains as follows.

**Definition 5.25.** [27] A set \( M \subset D \subseteq \mathbb{R}^n \) is a positively invariant set with respect to the dynamical system of Definition 5.24 if \( s(M) \subseteq M \) for all \( t \geq 0 \), where \( s(M) = \{ s_t(x) | x \in M \} \), with \( s_t(x) \) is given as in Definition 5.23.

Last, using Definition 5.26, we turn our focus to results on the stability of equilibrium points of the IVP of Definition 5.24, defined below.

**Definition 5.26.** [60] Consider the dynamical system of Definition 5.24. A point \( x^* \in D \) is said to be an equilibrium point of Equation (5.1) if \( f(t, x^*) = 0 \) for all \( t \in \mathbb{R} \).

To examine stability of equilibria, we can undertake a *Lyapunov stability analysis*, which studies the behaviour of solutions to the dynamical system about an equilibrium. An important result was first proposed by A. Lyapunov in 1892 [60], and prescribes conditions for the stability of an equilibrium point. Here we give the result for the 2-dimensional system

\[
\frac{dx}{dt} = f(x, y) \quad \text{and} \quad \frac{dy}{dt} = g(x, y), \quad (5.2)
\]
however, this result can be generalised for an N-dimensional system.

**Definition 5.27.** [1] Let $U$ be an open subset of $\mathbb{R}^2$ containing the origin. A real-valued $C^1$ function $V : U \rightarrow \mathbb{R}, [(x, y) \in U, V(x, y) \in \mathbb{R}]$ is said to be positive definite on the set $U$ if the following two conditions hold:

1. $V(0, 0) = 0$.

2. $V(x, y) > 0$ for all $(x, y) \in U$ with $(x, y) \neq (0, 0)$.

**Theorem 5.4.** [1] Let $(0, 0)$ be an equilibrium of the autonomous system given in Equation (5.2) and let $V$ be a positive definite $C^1$ function in a neighbourhood $U$ of the origin.

1. If $\frac{dV(x,y)}{dt} \leq 0$ for $(x, y) \in U \setminus \{(0,0)\}$, then $(0,0)$ is stable.

2. If $\frac{dV(x,y)}{dt} < 0$ for $(x, y) \in U \setminus \{(0,0)\}$, then $(0,0)$ is asymptotically stable.

3. If $\frac{dV(x,y)}{dt} > 0$ for $(x, y) \in U \setminus \{(0,0)\}$, then $(0,0)$ is unstable.

In Case 1, the function $V$ is a Lyapunov function and in Case 2 $V$ is a strict Lyapunov function.
Bibliography


[80] R. Sivakumar, N. Dimopoulos, and W. Lu. Analytical formulas for \(\sum_{i=1}^{n} \left\lfloor \frac{i}{p} \right\rfloor\) and \(\sum_{i=1}^{n} \left\lceil \frac{i}{p} \right\rceil\). *Pi Mu Epsilon Journal*, 10:458–460, 1997.


