Stochastic Modelling of Subcellular Biochemical Systems

Dissertation
zur
Erlangung des akademischen Grades
Doktor-Ingenieur (Dr.-Ing.)

Promotionsgebiet Systembiologie

Fakultät für Informatik und Elektrotechnik
Universität Rostock

vorgelegt von

Mukhtar Ullah, geb. am 16. April 1975 in Malakand, Pakistan

aus Rostock

Rostock, 6. März 2009

urn:nbn:de:gbv:28-diss2009-0128-8

Gutachter:

Prof. Olaf Wolkenhauer (Institut für Informatik, Universität Rostock)
Prof. Hanspeter Herzel (Institut für Theoretische Biologie, Humboldt Universität Berlin)
Prof. Volkmar Liebscher (Institut für Mathematik und Informatik, Universität Greifswald)

Tag der Promotionsverteidigung: 29. Juni 2009

This work is dedicated to minds that are
open to rational scientific reasoning and
are not pre-occupied with dogmas.

Abstract

While ordinary differential equations (ODEs) form the conceptual framework for modelling many cellular processes, specific situations demand stochastic models to capture the influence of noise. Motivated by that, we seek a stochastic framework based on Markov processes to represent intracellular processes. We review the formal relationships between different stochastic approaches referred to in the systems biology literature.

The most common formulation of stochastic models for biochemical networks is the chemical master equation (CME). While stochastic simulations are a practical way to realise the CME, analytical approximations offer more insight into the influence of randomness. Towards that end, the two-moment approximation (2MA) is a promising addition to the established analytical approaches including the chemical Langevin equation (CLE) and the related linear noise approximation (LNA). The 2MA approach directly tracks the mean and (co)variance which are coupled in general. This coupling is not obvious in CME and CLE and ignored by LNA and conventional ODE models.

We extend previous derivations of 2MA by allowing a) non-elementary reactions and b) relative concentrations. Often, several elementary reactions are approximated by a single step. Furthermore, practical situations often require the use of relative concentrations.

We investigate the applicability of the 2MA approach to the well established fission yeast cell cycle model. Our analytical model reproduces the clustering of cycle times observed in experiments. This is explained through multiple resettings of a protein called MPF, caused by the coupling between mean and (co)variance, near the G2/M transition.

Abstract in German (Zusammenfassung)

Während gewöhnliche Differentialgleichungen (ordinary differential equations - ODEs) den konzeptionellen Rahmen für die Modellierung vieler zellulärer Prozesse bilden, erfordern spezielle Situationen stochastische Modelle, um den Einfluss von Zufälligkeit mit einzubeziehen. Hierdurch motiviert suchen wir eine stochastische Formulierung System basierend auf Markov-Prozessen, welches innerzelluläre Prozesse repräsentiert. Wir diskutieren die formalen Beziehungen zwischen den verschiedenen stochastischen Ansätzen auf die, in der Literatur der Systembiologie Bezug genommen wird.

Die gebräuchlichste Form von stochastischen Modellen für biochemische Netzwerke ist die chemische Mastergleichung (chemical master equation – CME). Während stochastische Simulationen ein praktischer Weg sind, um die CME numerisch zu loesen, ermöglichen analytische Näherungen eine bessere Sicht auf den Einfluss von zufälligen Variationen. Für diesen Zweck ist die 'two-moment approximation' (2MA) eine vielversprechende Erweiterung zu den bewährten analytischen Ansätzen, inklusive der chemischen Langevin-Gleichung (chemical Langevin equation - CLE) und der verwandten 'linear noise approximation' (LNA). Der 2MA Ansatz beschreibt den Mittelwert und die (Ko-)Varianz, die miteinander gekoppelt sein können. Diese Koppelung ist ersichtlich in der CME und der CLE und wird bei der LNA und den konventionellen und konventionellen ODE-Modellen ignoriert.

Wir erweitern bisherige Herleitungen des 2MA mit Bezug auf a) nicht-elementare Reaktionen und b) relative Konzentrationen. Dies ist dadurch motiviert das oft mehrere elementare Reaktionen zu einem einzelnen Schritt zusammengefasst werden und in praktischen Situationen die Anwendung relativer Konzentrationen von Bedeutung ist. Desweiteren erfordern praktische Situationen die Anwendung relativer Konzentrationen.

Wir untersuchen die Anwendbarkeit des 2MA Ansatzes am Beispiel eines Zellzyklusmodells. Unser analytisches Modell spiegelt die Gruppierung der in Experimenten beobachteten Zykluszeiten wieder. Dieses kann durch das mehrfache Zurücksetzen des Proteins MPF erklärt werden, hervorgerufen durch die Kuppelung zwischen Mittelwert und (Ko-)Varianz nahe des G2/M Übergangs.

Acknowledgments

I am grateful to my supervisor Olaf Wolkenhauer for his consistant guidance and support during my research.

I acknowledge the support of the

1. U.K. Department for the Environment, Food and Rural Affairs (DEFRA) [Oct 03 to Sep 05]
2. University of Rostock [Oct 05 to Sep 08]
3. Deutsche Forschungsgemeinschaft (DFG) through grant (WO 991/3-1). [Oct 08 to date]

I thank the following people for their willingness to discuss my work during my doctoral research: Thomas Millat (University of Rostock), Hanspeter Herzel (Humboldt University, Berlin), Akos Sveiczer (Budapest University of Technology and Economics), Kevin Burrage (University of Queensland, Brisbane), Allan Muir (Llanfynwydd, Wales) and Johan Elf (Uppsala University, Sweden).

I would like to add my gratitude to other members of the systems biology and bioinformatics group, in particular Ulf Schmitz.

Finally, I would like to thank my wife Sabahat Ahmad, who gave me the strength and love I needed during the course of this work.

Errata

Page 2, line 18 : reportd → reported

Page 2, last paragraph: Fluctuation-effected → Fluctuation-affected

---

Page 4, 30th line: the the → the

Page 30, 6th line: biomolecular → bimolecular

Page 37, 19th line: $\mu_{N,i} \rightarrow \langle N_i \rangle$

Page 38, 5th line: the the → the

Page 47, line 9: weakend → weakened

Page 52, Table 4.4 header row: $f_i \rightarrow \tilde{f}_i$

Page 58, Table 4.6:

Case	μ_{CT}	σ_{CT}	CV_{CT}	μ_{dM}	σ_{dM}	CV_{dM}	μ_{BM}	σ_{BM}
(1)	131	47	0.358	2.22	0.45	0.203	1.21	0.24
(2)	138.8	12.4	0.09	1.59	0.058	0.0362	3.18	0.101
(3)	138.8	17.6	0.127	1.62	0.093	0.0576	3.25	0.178
(4)	138.8	23.9	0.172	1.66	0.12	0.0721	3.32	0.231

→

Case	μ_{CT}	σ_{CT}	CV_{CT}	μ_{DM}	σ_{DM}	CV_{DM}	μ_{BM}	σ_{BM}
(1)	131	47	0.358	2.22	0.45	0.203	1.21	0.24
(2)	138.8	12.4	0.09	3.18	0.101	0.0319	1.59	0.0575
(3)	138.8	17.6	0.127	3.25	0.178	0.055	1.623	0.0934
(4)	138.8	23.9	0.172	3.32	0.231	0.0697	1.657	0.12

Page 58, 13th line of the last paragraph:

The mean BM is much larger than the experimental BM

→ The mean values for both BM and DM are larger than the corresponding experimental values

---

Page 71, last reference: , 2006.

→ . In *Computational Intelligence and Bioinformatics*, pages 786–791. Springer Berlin, 2006.

Page 73, first publication heading:

Revised manuscript submitted to JTB ...

→ *Journal of Theoritical Biology*, doi:10.1016/j.jtbi.2009.05.022

Contents

Acronyms

ODE	ordinary differential equation
CME	chemical master equation
CLE	chemical Langevin's equation
SSA	stochastic simulation algorithm
LNA	linear noise approximation
FPE	Fokker-Planck equation
CKE	Chapman-Kolmogorov equation
dCKE	differential Chapman-Kolmogorov equation
2MA	two-moment approximation
DA	degree of advancement
CDF	cumulative distribution function
CCDF	complementary CDF
PDF	probability density function
MPF	metaphase promoting factor
CT	cycle time
BM	birth mass
DM	division mass
CV	coefficient of variation
NSR	noise-to-signal ratio
xNSR	cross noise-to-signal ratio
std	standard deviation

Notation

General notes

- Symbols of the form $Q(t)$ represent time-dependent quantities.
- Every integer-indexed quantity q_i is also the ith element of a corresponding vector q of appropriate dimensions. Similarly, every integer-pair indexed quantity Q_{ij} is also the (i,j)th element of a corresponding matrix Q of appropriate dimensions.
- For a matrix S, the transpose is written S^T, the ith row is denoted $S_{i\bullet}$ and the jth column is denoted by $S_{\bullet j}$.
- Any symbol denoting a random/stochastic variable/process has to be in capital. Once defined, the corresponding symbol in small represents a sample/realisation of the variable/process. The time-dependent version of the same symbol in small represents a deterministic approximation of the corresponding stochastic process. Thus if $N(t)$ is a stochastic process, n is a typical sample of $N(t)$ and $n(t)$ is a deterministic approximation of $N(t)$.
- Symbols denoting operators such as probability are typset upright, e.g. $\Pr[\cdot]$.
- The notation $f(\cdots, n, \cdots)$ is short-hand for $f(\cdots, n_1, \cdots, n_s, \cdots)$ whenever an s-vector n appears as an argument.

List of symbols

X_i	ith chemical species/component
$\varnothing$	null species
s	number of chemical components
R_j	jth reaction channel
r	number of reaction channels
$N_i(t)$	copy number
$N^c(t)$	continuous approximation of $N(t)$

$X_i(t)$	concentration
N_{A}	Avogadro's constant
V	volume
Ω	system-size when scalar; scaling parameters when vector
S	stoichiometry matrix
$Z_j(t)$	number of R_j occurrences during $[0, t]$
$\Pr[\cdot]$	probability measure
$P(n, t)$	state probability
$P^c(n, t)$	continuous approximation of $P(n, t)$
$P(n\|m, t)$	transition probability
$v_j(x)$	reaction rate
$a_j(n)$	reaction propensity
$\tilde{a}_j(x)$	stochastic reaction rate
$a_0(n)$	sum of $a_j(n)$ over all reaction channels (exit rate)
$\mathbb{E}_j$	a step operator defined by $\mathbb{E}_j f(n) = f(n - S_{\cdot j})$
$\langle Y(t) \rangle$	mean/expectation of the process $Y(t)$
$\left\langle Y, Y^T \right\rangle$	covariance matrix
$\langle Y_i, Y_k \rangle$	(i, k)th element of the above covariance matrix
$\mathrm{Var}[\cdot]$	variance
$\mu_i(t)$	mean concentration
$\sigma_{ik}(t)$	pair-wise concentration covariance
$\zeta_{ii}(t)$	NSR
$\zeta_{ik}(t)$	xNSR

$f_i(n)$	copy-number flux
$B(n)$	copy-number diffusion matrix
$\tilde{f}(x)$	concentration flux
$\tilde{B}(x)$	concentration diffusion matrix
$A \odot B$	Hadamard product - element-wise product
$A : B$	Frobenius inner product - sum of elements of $A \odot B$
$\mathcal{P}_j$	Poisson random variable
$\mathcal{N}_j$	standard normal random variable
$\mathcal{W}_j$	standard brownian motion, or Wiener process
$\mathcal{O}(x)$	first neglected order with respect to x in an expansion
$o(x)$	terms vanishing faster than x, as the latter approaches zero

List of Figures

List of Tables

1

Introduction

1.1 Why stochastic modelling?

At a coarse level, cellular functions are largely determined by spatio-temporal changes in the abundance of molecular components. At a finer level, cellular events are triggered by discrete and random encounters of molecules [PE06]. This suggests a deterministic modelling approach at the coarse level (cell function) and a stochastic one at the finer level (gene regulation) [RWA02, Pau04, KEBC05, RO05, PvO05, Man07, ABRB$^+$08, LPB$^+$06, Pas07, BKvO05]. However, stochastic modelling is necessary when noise propagation from processes at the fine level changes cellular behaviour at the coarse level.

Stochasticity is not limited to low copy numbers. The binding and dissociation events during transcription initiation are the result of random encounters between molecules [KEBC05]. If molecules are present in large numbers and the molecular events occur frequently, the randomness would cancel out (both within a single cell and from cell to cell) and the average cellular behaviour could be described by a deterministic model. However, many subcellular processes, including gene expression, are characterised by infrequent (rare) molecular events involving small copy numbers of molecules [KEBC05, PE06]. Most proteins in metabolic pathways and signalling networks, realising cell functions, are present in the range 10-1000 copies per cell [BPE00, LKM07, Pau05]. For such moderate/large copy numbers, noise can be significant when the system dynamics are driven towards critical points in cellular systems which operate far from equilibrium [EE03, TJD05, ZYDQ06]. The significance of noise in such systems has been demonstrated for microtubule formation [DL93], ultrasensitive modification and demodification reactions [BPE00], plasmid copy number control [PE01], limit cycle attractor [Qia02], noise-induced oscillations near a macroscopic Hopf bifurcation [VKBL02], and intracellular metabolite concentrations [EPBE03].

Noise has a role at all levels of cell function. Noise, when undesired, may be suppressed by the network (e.g. through negative feedback) for robust behaviour [SK04, TvO02, FHG$^+$04, MA04, RWA02, PE00]. However, all noise may not be rejected and some noise may even be amplified from process to process, and ultimately influencing the phenotypic behaviour of the cell [HB08, LP06, PvO05, BKvO05, SU08]. Noise may even be exploited by the network to generate desired variability (phenotypic and cell-type diversification) [Blo06, CW06, HPDC00, RWA02, YUIS07]. Noise from gene expression can induce new dynamics including signal amplification [SPA05], enhanced sensitivity (stochastic focusing)

[PBE00, PvO05], bistability (switching between states) and oscillations [FX01, AMP07, OTL$^+$04, ADKC07, LL08], stabilization of a deterministically unstable state [TGOS08] and even discreteness-induced switching of catalytic reaction networks [TK07]. These are both quantitatively and qualitatively different from what is predicted or possible deterministically.

In the rest of the present section, we illustrate the need for stochastic modelling by selecting a few important aspects of biochemical reaction networks.

Identifiability: In the isomerisation reaction $X_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X_2$, proteins are converted back and forth between the inactive form X_1 and the active form X_2 such that the total number n^{tot} of protein molecules remains constant. When treated deterministically, the number n of proteins in the inactive form varies continuously with time according to the ODE, to be derived in the next chapter,

$$\frac{\mathrm{d}n}{\mathrm{d}\tau} = \frac{k_2 n^{\text{tot}}}{(k_1 + k_2)} - n,$$

where k_1 and k_2 are the respective rate constants of the activation and inactivation, and τ denotes a non-dimensional time variable. Here we see that $n(\tau)$ depends on the fraction $\hat{k} = {k_2}/{(k_1+k_2)}$ but not on the particular values of k_1 and k_2. In other words, experimental data on protein copy numbers can only provide information about the fraction $\hat{k}$, and not on the particular values of k_1 and k_2 separately. This issue of *identifiability* is reportd in [Wil06, Wil09]. The problem here is that changes in the protein copy numbers are discrete and random, rather than continuous and deterministic. We will learn in Chapter 3 that the variance of n satisfies an ODE which involves the difference $k_1 - k_2$ between the two parameters, in addition to the fraction $\hat{k}$. Thus experimental data on fluctuations, combined with the experimental data on the average protein copy numbers, would give information about both k_1 and k_2 separately.

Extinction: Due to the continuous treatment of the copy number, $n(t)$ in the above example can never become zero for non-zero rate constants. However, when treated stochastically, discrete changes admit questions to be asked about the probability of *extinction*, $n(t) = 0$, and about the average time until the first extinction.

Fluctuation-effected mean: In the isomerisation example, the mean (copy number) of the stochastic model was the same as the solution of the corresponding deterministic model. However, we will learn in Chapter 3 that this is not true in general. For system containing bimolecular reactions, the mean is also influenced by the fluctuations. In some systems, the mean of the stochastic model can be considerably larger than the deterministic prediction, and can lead to enhanced sensitivity of the network, known as *stochastic focusing* [PBE00, PvO05].

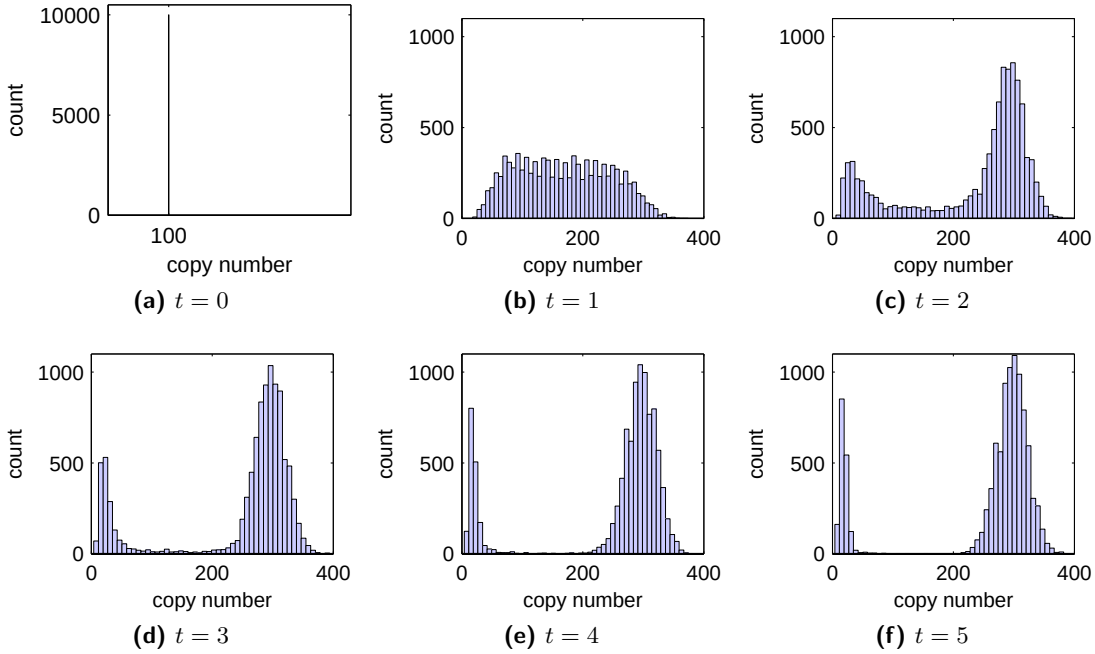

Figure 1.1 The time development of the histogram for the Schlögl reaction.

Bistability: A bistable system has two stable steady states separated by an unstable steady state. In a deterministic framework, such a system settles to that steady state whose basin of attraction is nearer to the initial condition. In a stochastic framework, however, the behaviour is more complex: either steady state may be reached in different realisations regardless of the initial condition. This behaviour is referred to as *stochastic switching* [UIYS06, GUV07], illustrated in Figure 1.1 for the Schlögl reaction, to be discussed in the following two chapters. The time varying histogram, which was obtained from 10000 realisations, is unimodal initially and has a bi-modal pattern at the end.

To study a cell that can be in different states (e.g. apoptosis, cell differentiation), single cell technologies are necessary. Averaging over ensembles of cells, as done in a western blot, does not allow to disinguish between states. Using single cell technologies, such as microscopy, a sample generated from a collection of cells under the same condition has proportions of cells in each state. Stochastic approach is necessary for capturing the variability in these experimental observations.

The notion of noise: The term *noise* can be confusing because it is not uniquely defined for all systems. Similarly the classification of noise (e.g. internal/external) can have different meanings for different system. However, noise and its various kinds in gene expression have been clearly defined in [KEBC05, RO05, Pau05, DMK$^+$06, BKvO05]. Following [RO05], noise in gene expression refers to the stochastic variation of a (expressed) protein concentration within isogenic cells having the same history and conditions (environment). Placing two gene reporters in the same cell and quantifying their gene expression (by the abundance of their target proteins) allows the following categorisation of noise (see

Figure 2 in [RO05]). *Intrinsic noise* arises from sources that create differences (in the gene expression) between the two reporters in the same cell, and *extrinsic noise* arises from sources that have equal effect on the two reporters in the same cell but create differences between two cells. Stochastic events during gene expression would then emerge as intrinsic noise whereas differences between cells will appear as extrinsic noise. Extrinsic noise can be *global* when fluctuations in basic reaction rates affect expressions of all genes, or it can be *pathway-specific*. It is important to realise that extrinsic noise can be theoretically isolated from the system but intrinsic noise is the very essence (discrete nature) of the underlying molecular events and cannot be separated (even hypothetically) from the system. Finally, we like to add that the word "noise" has often negative associations as something undesirable, something that should be removed or avoided. In biology, noise can also have a role and "randomness" may be a better word. In this text the word 'noise' is used, with the understanding that it may well be something desirable.

1.2 Established stochastic approaches

The most common formulation of stochastic models for biochemical networks is the chemical master equation (CME). While stochastic simulations [TSB04, Pah08] are a practical way to realise the CME, analytical approximations offer more insights into the influence of noise on cell function. Formally, the CME is a continuous-time discrete-state Markov process [Sin53, Gil77, Kam07a]. For gaining intuitive insight and a quick characterisation of fluctuations in biochemical networks, the CME is usually approximated analytically in different ways [Kam07a, Gou06], including the frequently used the chemical Langevin approach [Gil00, Kam07b, Ste04, ZHCN07], the linear noise approximation (LNA) [EE03, HJ04, SI05, SIK06] and the two-moment approximation (2MA) [Gou07, GUV07, FLH07].

Of the analytical approaches mentioned above, we here focus on the 2MA approach because of its representation of the coupling between the mean and (co)variance. The traditional Langevin approach is based on the assumption that the time-rate of abundance (copy number or concentration) or the flux of a component can be decomposed into a deterministic flux and a Langevin noise term, which is a Gaussian (white noise) process with zero mean and amplitude determined by the the dynamics of the system. This separation of noise from the system dynamics may be a reasonable assumption for *external noise* that arises from the interaction of the system with other systems (like the environment), but cannot be assumed for internal noise that arises from within the system [KEBC05, RO05, Pau05, BKvO05, DMK$^+$06, SOS08]. As categorically discussed in [Kam07b], internal noise is not something that can be isolated from the system because it results from the discrete nature of the underlying molecular events. Any noise term in the model must be derived from the system dynamics and cannot be presupposed in an *ad hoc* manner. However the chemical Langevin equation (CLE) does not suffer from the above criticism because Gillespie [Gil00] derived it from the CME description. The CLE allows much faster simulations compared to the exact stochastic simulation algorithm (SSA) [Gil77] and its variants. The CLE is a stochastic differential equation (dealing directly with random variables rather than moments) and has no direct way of representing the mean and (co)variance and the coupling between the two. That does

not imply that CLE ignores the coupling like the LNA which has the same mean as the solution of the deterministic model.

The merits of the 2MA compared to alternative approximations have been discussed in [GUV07, Gou07, Tan08]. In [FLH07], the 2MA is developed as an approximation of the master equation for a generic Markov process. In [GUV07], the 2MA framework is developed under the name "mass fluctuation kinetics" for biochemical networks composed of elementary reactions. The authors demonstrate that the 2MA can reveal new behaviour like stochastic focusing and bistability. Another instance of the 2MA is proposed in [Gou06, Gou07] under the names "mean-field approximation" and "statistical chemical kinetics". Again, the authors assume elementary reactions so that the propensity function is at most quadratic in concentrations. The authors evaluate the accuracy of the 2MA against the alternatives (such as LNA) for a few toy models. The derivation of the 2-MA for more general systems with non-elementary reactions is one motivation for the present paper.

The 2MA approaches referred to above assume absolute concentrations (copy number divided by some fixed system size parameter). In systems biology, however, models often use relative concentrations that have arbitrary units [NPCT01, NCT05, TCNN02, CNBC$^+$06]. In general, the concentration of each component in the system may have been obtained by a different scaling parameter, rather than using a global system size. For such models, the above mentioned approaches need modification. This was another motivation for our derivation in this paper.

1.3 Research objectives

While most of the literature in systems biology focuses on numerical solutions and stochastic simulations, the focus of the present work here is on analytical approaches. In this text we develop a compact form of the 2MA equations - a system of ODEs for the dynamics of the mean and (co)variance of the continuous-time discrete-state Markov process that models a biochemical reaction system by the CME. This is an extension of previous derivations, taking into account arbitrary concentrations and non-elementary reactions. The compact form, obtained by careful selection of notation, of our derivation allows for an easy interpretation. Using these analytical results, we develop our 2MA model of the fission yeast cell cycle which has two sets of ODEs: one set for the mean protein concentrations and the other set for concentration (co)variances. Numerical simulations of our model show a considerably different behaviour. Especially, for the *wee1*$^-$ *cdc25*Δ mutant (hereafter referred simply as double-mutant), the timings of S-phase and M-phase are visibly different from those obtained for a deterministic model because of the oscillatory behaviour of the key regulator. Since the 2MA is only an approximation, we investigate its validity by comparing the statistics computed from the 2MA model with experimental data. In summary, the research objectives of the present work are to:

- Present the stochastic description in a notation appropriate for system biology.
- Review various approximations to a stochastic process and clarify relationships between them.

- Seek an analytical approach that bridges the gap between the deterministic and stochastic approaches by combining the intuition of derministic models with the represention of noise and variability.
- Generalise the two-moment approximation to allow for non-elementary reactions and relative concentrations.
- Apply the above theoretical approach to a real world application and compare the model predictions with experimental data (here cell cycle).
- Demonstrate the value of the 2MA approach in comparison to the deterministic approach using ODEs and the alternative stochastic approaches.

1.4 Outline of the text

The rest of the present work is organised as follows.

Chapter 2: We present the stochastic framework for modelling subcellular biochemical systems. In particular, we make an effort to show how the notion of propensity, the chemical master equation and the stochastic simulation algorithm arise as consequences of the Markov property. This connection is not obvious from the relevant literature in systems biology. Moreover, we review various analytical approximations of the chemical master equation. The chapter is concluded with a sketch of interrelationships between various stochastic approaches.

Chapter 3: This chapter develops a compact form of the 2MA equations - a system of ODEs for the dynamics of the mean and (co)variance of the continuous-time discrete-state Markov process that models a biochemical reaction system by the CME. This is an extension of previous derivations, taking into account relative concentrations and non-elementary reactions. The compact form, obtained by careful selection of notation, allows for an easy interpretation.

Chapter 4: This chapter takes the Tyson-Novák model for the fission yeast cell cycle as a case study. This deterministic model is a practical example using non-elementary reactions and relative concentrations, the two central features of our extended 2MA approach. This will allow us to investigate the price of higher-order truncations by comparing the simulated cycle time statistics with experiments.

Conclusions: Here we summarise our key arguments, results and give suggestions for further work.

Publications: A list of publications that have arisen from this work, including those whose content has not been discussed in the present text in order to keep it concise.

2

Stochastic modelling

In this chapter, we present a stochastic framework for modelling subcellular biochemical systems. In particular, we make an effort to show how the notion of propensity, the chemical master equation (CME) and the stochastic simulation algorithm arise as consequences of the Markov property. This connection is not obvious from the relevant literature in systems biology. We review various analytical approximations of the CME, leaving out stochastic simulation approaches reviewed in [TSB04, Pah08]. Moreover, we sketch interrelationships between various stochastic approaches. The books by [PP01] and [Wil06] inspired this chapter.

2.1 Chemical reactions and species

Imagine molecules of s chemical species homogeneously distributed in a compartment of constant volume V at thermal equilibrium and interacting through r irreversible (unidirectional) reaction channels. A reaction channel is usually assumed to be a single step in which case it is called *elementary*. However, it may represent a simplification of multiple elementary steps into a single step. Any reversible (bidirectional) reaction can be listed as two irreversible reactions. We symbolise the ith species with X_i and the jth reaction channel with R_j. The abundance of X_i present in the system at time t can be described by the copy number $N_i(t)$ or the concentration

$$X_i(t) = \frac{N_i(t)}{\Omega} \,.$$

The scaling parameter Ω is called the *system size*. For molar concentrations the system size is chosen as $\Omega = N_\mathrm{A} V$ where N_A is the Avogadro's constant. For relative concentrations the system size is some fixed copy number. Take the isomerisation reaction as an example where proteins are converted back and forth between the inactive form X_1 and the active form X_2 such that the total number n^tot of protein molecules remains constant. The relative concentrations in this example are the fractions,

$$X_1(t) = \frac{N_1(t)}{n^\mathrm{tot}} \quad \text{and} \quad X_2(t) = \frac{n^\mathrm{tot} - N_1(t)}{n^\mathrm{tot}}$$

of proteins in the inactive and active form, respectively. It is sometimes more appropriate to use a different scaling parameter Ω_i for each component i. This will of concern to us in the following chapter. In this chapter, we stick to the simpler case.

The reaction channel R_j will be represented by the general scheme

$$R_j: \quad \underline{S}_{1j}\mathrm{X}_1 + \cdots + \underline{S}_{sj}\mathrm{X}_s \xrightarrow{k_j} \bar{S}_{1j}\mathrm{X}_1 + \cdots + \bar{S}_{sj}\mathrm{X}_s\,. \tag{2.1}$$

The coefficient $\underline{S}_{ij}$ (on the left) represents the participation of X_i as reactant and $\bar{S}_{ij}$ (on the right) is the corresponding participation as product. These coefficients are called *stoichiometries* or *stoichiometric coefficients*. The rate constant, or coefficient, k_j, written over the reaction arrow will be explained later. The progress of channel R_j is quantified by the *degree of advancement* (DA) $Z_j(t)$ defined as the number of occurrences of R_j during the time interval $[0,t]$. One occurrence of R_j changes the copy number of X_i by $S_{ij} = \bar{S}_{ij} - \underline{S}_{ij}$, the (i,j)th element of the *stoichiometric matrix* S. During the time interval $[0,t]$, the change in the copy number of X_i contributed by R_j is $S_{ij}Z_j(t)$. The total change in the copy number is the sum of contributions from all reactions:

$$N_i(t) = N_i(0) + \sum_{j=1}^{r} S_{ij}Z_j(t)\,. \tag{2.2}$$

Thus changes in copy numbers are determined by stoichiometries and DAs. Following the usual vector notation, we write $N(t)$ for the $s\times 1$-vector of copy numbers, $X(t)$ for the $s\times 1$-vector of concentrations and $Z(t)$ for the $r\times 1$-vector of DAs. The above conservation relation can be written in the vector notation:

$$N(t) = N(0) + S\,Z(t)\,. \tag{2.3}$$

Dividing by Ω gives the corresponding relation in concentrations:

$$X(t) = X(0) + \frac{S\,Z(t)}{\Omega}\,. \tag{2.4}$$

The copy number $N(t)$, the concentration $X(t)$ and the DA $Z(t)$ are alternative ways to describe our system. Description in terms of these *macroscopic variables* is done in the hope that they approximately satisfy an autonomous set of deterministic equations. Two problems stand in making such an effort. First, the reactions are discrete events in time which means that the copy numbers do not vary continuously with time. Secondly, the occurrence time of a reaction is a random quantity because it is determined by a large number of microscopic factors (e.g. positions and momenta of the molecules involved). Therefore, the deterministic description needs a few simplifying assumptions. Alternatively the macroscopic variables are formulated as stochastic processes. Such a stochastic description in terms of macroscopic variables is called *mesoscopic*.

Throughout this text, we will use the following academic examples. They are chosen to demonstrate different ideas and methods in the discussion.

Isomerisation: Consider a protein that can exist in two different forms, an *inactive* form X_1 and an *active* form X_2. The protein changes between the two forms by the reversible isomerisation reaction

$$\mathrm{X}_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{X}_2 \tag{2.5}$$

composed of an inactivation (forward) channel with rate constant k_1 and an activation (reverse) channel with rate constant k_2. The reaction scheme (2.5) also represents the opening and closing of an ion-channel and similar systems with two-state conformational change. This example was used in the introduction to illustrate ideas of identifiability and species extinction.

Lotka-Volterra model: Consider a system consisting of two interacting species: X_1 and X_2. The species can either be animals (X_1: prey, X_2: predator), chemical species or any interacting entities of two kinds. A large amount of a substance A is available for X_1 which reproduces immediately after consuming one unit A. An encounter between the two species results in the disappearance of X_1 and the replication of X_2. This is the only way X_1 dies (degrades) whereas X_2 has a natural death (degradation). The system can be represented by the following scheme

$$\left.\begin{aligned} X_1 + A &\xrightarrow{k_1} 2X_1 \\ X_1 + X_2 &\xrightarrow{k_2} 2X_2 \\ X_2 &\xrightarrow{k_3} \varnothing \end{aligned}\right\} \tag{2.6}$$

with rate constants k_1, k_2 and k_3. The symbol $\varnothing$ represents the dead (degraded) form of X_2. In general, the "null species" represents any species not included in the model. The substance A is constantly replenished so that the copy number n_A remains constant. This system was first investigated by Lotka and Volterra [Lot20, Vol26]. Here it serves the purpose of a simple system containing a bimolecular reaction and the resulting influence of (co)variance on the mean.

Enzyme kinetic reaction: The enzyme kinetic reaction

$$E + S \longrightarrow E + P$$

can be decomposed into a set of three elementary reactions:

$$\left.\begin{aligned} E + S &\xrightarrow{k_1} ES \\ ES &\xrightarrow{k_2} E + S \\ ES &\xrightarrow{k_3} E + P \end{aligned}\right\} \tag{2.7}$$

Here the enzyme E catalyses a substrate S into a product P that involves an intermediary complex ES. We include this example because this type of reaction appears frequently in the literature. It also serves the purpose of a simple system containing a bimolecular reaction and how a mass conservation leads to a simplified model.

Schlögl model: An autocatalytic, trimolecular reaction scheme, first proposed by [Sch72]

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

can be decomposed into

$$\left.\begin{aligned} \mathrm{A} + 2\mathrm{X} &\xrightarrow{k_1} 3\mathrm{X} \\ 3\mathrm{X} &\xrightarrow{k_2} \mathrm{A} + 2\mathrm{X} \\ \mathrm{B} &\xrightarrow{k_3} \mathrm{X} \\ \mathrm{X} &\xrightarrow{k_4} \mathrm{B} \end{aligned}\right\} \tag{2.8}$$

Here the concentrations A and B are kept constant (buffered). This example, mentioned in the introduction, serves to illustrate the need for a stochastic approach to model systems with bistability.

2.2 Deterministic description

Suppose that the reactions occur so frequently that the DA $Z(t)$ can be approximated by a continuous quantity $z(t)$. This assumption requires that a large number of reactant molecules are freely available (no crowding) in a large volume so that they can react easily. It also requires that the energy and orientation of reactant molecules favour the reaction, a fact summarised in a rate constant. Large numbers of molecules also mean that a change resulted from a single occurrence of a reaction is relatively small. That means that the copy number $N(t)$ can be approximated by a continuous quantity $n(t)$. The concentration $X(t)$ is similarly approximated by a continuous quantity $x(t)$. In deterministic description, equations (2.3) and (2.4) respectively translate to

$$n(t) = n(0) + S\, z(t)\,. \tag{2.9}$$

and

$$x(t) = x(0) + \frac{S\, z(t)}{\Omega}\,. \tag{2.10}$$

Taking the time-derivatives we arrive at the deterministic *chemical kinetic equations*:

$$\dot{n}(t) = \Omega S\, v\left(\frac{n(t)}{\Omega}\right) \quad \text{and} \quad \dot{x}(t) = S\, v\left(x(t)\right),$$

where $v = \dot{z}/\Omega$ is the *reaction rate* vector whose ith element v_j is the rate of R_j. The reaction rate v_j is the number of occurrences of R_j per unit time divided by the system size. The notation $v\left(x(t)\right)$ is based on the assumption that the reaction rate depends only on the concentrations of the reactants. This is a realistic assumption in many reactions at constant temperature. In general, the reaction rate can depend on temperature, pressure, and the concentrations or partial pressures of the substances in the system.

The functional form $v_j(\cdot)$ of the rate of R_j is called the *rate law*. There is a large class of chemical reactions in which the reaction rate is proportional to the concentration of each reactant raised to some power:

$$v_j(x) = k_j \prod_{i=1}^{s} x_i^{g_{ij}}, \tag{2.11}$$

which is called *a rate law with definite orders* [Mor08]. The rate constant k_j summarises factors such as enough energy and proper orientation of the reactant molecules to have an encounter leading to the reaction. The exponent g_{ij} is the order with respect to the species X_i. The sum of orders for a particular reaction channel is the overall order. For elementary reactions, the orders g_{ij} are the same as the reactant stoichiometries $\underline{S}_{ij}$:

$$v_j(x) = k_j \prod_{i=1}^{s} x_i^{\underline{S}_{ij}} . \tag{2.12}$$

This rate law is called *mass action kinetics* [HS96] and is justified by collision theory and transition state theory [Wri04, Hou07, Mor08]. Reactions that cannot be described by rate laws like (2.11) are said not to have a definite order. For such a reaction, the rate law depends on the assumptions involved in the approximation of the constituent reaction channels. Examples of rate laws for the reactions are Michaelis-Menten kinetics, Hill kinetics and competitive inhibition [Fel97, CB04, HS96].

Isomerisation revisited: The reaction scheme (2.5) depicts the reversible conversion of protein between two forms. Suppose there are n^{tot} copies of this protein in a container, $n(t)$ of them being in the inactive form X_1 at time t. The forward channel proceeds at a rate $\dot{z}_1 = k_1 n$ and the reverse channel does at a rate $\dot{z}_2 = (n^{\mathrm{tot}} - n)k_2$. The rate equation therefore reads

$$\dot{n} = -\dot{z}_1 + \dot{z}_2 = k_2 n^{\mathrm{tot}} - (k_1 + k_2)n . \tag{2.13}$$

With a non-dimensional, $\tau = (k_1 + k_2)t$, the ODE takes the form

$$\frac{\mathrm{d}n}{\mathrm{d}\tau} = \frac{k_2 n^{\mathrm{tot}}}{(k_1 + k_2)} - n .$$

Lotka-Volterra model revisited: The reaction scheme (2.6) depicts the predator-prey interactions. Let $n_1(t)$ and $n_2(t)$ denote the number of preys and predators, respectively. The number of food units is assumed to be so large that it is not changed by consumption during the time scale of our interest. The reaction rates according to the mass action kinetics follow from (2.12) to be

$$v_1 = k_1 n_{\mathrm{A}} n_1, \quad v_2 = k_2 n_1 n_2, \quad v_3 = k_3 n_2 .$$

The rate equations are then given by

$$\left.\begin{aligned} \frac{\mathrm{d}n_1}{\mathrm{d}t} &= v_1 - v_2 = k_1 n_{\mathrm{A}} n_1 - k_2 n_1 n_2, \\ \frac{\mathrm{d}n_2}{\mathrm{d}t} &= v_2 - v_3 = k_2 n_1 n_2 - k_3 n_2 . \end{aligned}\right\} \tag{2.14}$$

Enzyme kinetic reaction revisited: For the enzyme kinetic reaction (2.7), write $x_{\mathrm{E}}(t)$, $x_{\mathrm{S}}(t)$, $x_{\mathrm{ES}}(t)$ and $x_{\mathrm{P}}(t)$ for the respective the time-dependent molar concentrations of E, S, ES and P. The solution has to respect two conservation laws

$$x_{\mathrm{E}}(t) + x_{\mathrm{ES}}(t) = x_{\mathrm{E}}^{\mathrm{tot}}, \quad \text{and} \quad x_{\mathrm{S}}(t) + x_{\mathrm{ES}}(t) + x_{\mathrm{P}}(t) = x_{\mathrm{S}}^{\mathrm{tot}}$$

where $x_{\mathrm{E}}^{\mathrm{tot}}$ and $x_{\mathrm{S}}^{\mathrm{tot}}$ are the total concentrations of the enzyme and substrate, respectively. The channel-wise mass action kinetic law for the reaction scheme (2.7) can be written as:

$$v_j(x_{\mathrm{S}}, x_{\mathrm{ES}}) = \begin{cases} \left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} & \text{if } j = 1, \\ k_2 x_{\mathrm{ES}} & \text{if } j = 2, \\ k_3 x_{\mathrm{ES}} & \text{if } j = 3\,. \end{cases}$$

The concentrations evolve according to the following set of nonlinear coupled ODEs

$$\left.\begin{aligned} \frac{\mathrm{d}x_{\mathrm{S}}}{\mathrm{d}t} &= v_2 - v_1 = k_2 x_{\mathrm{ES}} - \left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}}, \\ \frac{\mathrm{d}x_{\mathrm{ES}}}{\mathrm{d}t} &= v_1 - v_2 - v_3 = \left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} - (k_2 + k_3)\, x_{\mathrm{ES}}\,. \end{aligned}\right\} \tag{2.15}$$

For the enzyme kinetic reaction (2.8) the Michaelis-Menten rate law takes the form

$$v\left(x_S\right) = \frac{V_{\max} x_{\mathrm{S}}}{K_{\mathrm{M}} + x_{\mathrm{S}}}$$

where $V_{\max} = k_2 x_{\mathrm{E}}^{\mathrm{tot}}$ and $K_{\mathrm{M}} = (k_1 + k_3)/k_2$. The concentrations then evolve according to the ODE

$$-\frac{\mathrm{d}x_{\mathrm{S}}}{\mathrm{d}t} = \frac{\mathrm{d}x_{\mathrm{P}}}{\mathrm{d}t} = v\left(x_{\mathrm{S}}\right)\,.$$

Schlögl model revisited: For the enzyme kinetic reaction (2.8), write x_{A} and x_{B} for the constant respective concentrations of chemicals A and B, and $x(t)$ for the time-dependent concentration of chemical X. The reaction rates according to the mass action kinetics follow from (2.12) to be

$$v_1 = k_1 x_{\mathrm{A}} x^2, \quad v_2 = k_2 x^3, \quad v_3 = k_3 x_{\mathrm{B}}, \quad v_4 = k_4 x\,.$$

The deterministic ODE then follows

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v_1 - v_2 + v_3 - v_4 = k_1 x_{\mathrm{A}} x^2 - k_2 x^3 + k_3 x_{\mathrm{B}} - k_4 x\,. \tag{2.16}$$

2.3 Stochastic mesoscopic description

The validity of macroscopic approaches for description of the averages is limited because the average of a nonlinear function is generally not the same as the function of the average. This was first demonstrated for bimolecular reactions in [Rén53].

Since the occurrence of reactions involve discrete events at the microscopic level, it is impossible to deterministically predict the progress of reactions in terms of the macroscopic variables such as $N(t)$ and $Z(t)$. To account for this uncertainty, one of the macroscopic variables $N(t), Z(t), X(t)$ is formulated as a stochastic process. Choosing the copy number $N(t)$, a sample value n of the process is the state of our biochemical system under consideration.

How does the process $N(t)$ evolve in time? Starting at time $t = 0$ from some initial state $N(0)$, every sample path of $N(t)$ chain remains in state $X(0)$ for a random amount of time W_1 until the occurrence of a reaction takes the process jumps to a new state $X(W_1)$; it remains in state $X(W_1)$ for another random amount of time W_2 until the occurrence of another reaction takes the process to a new state $X(W_1 + W_2)$, and so on. In other words, $N(t)$ is a *jump process*.

The stochastic process $N(t)$ is characterised by a collection of state probabilities and transition probabilities. The state probability

$$P(n,t) = \Pr\Big[N(t) = n\Big],$$

is the probability that the process $N(t)$ is in state n at time t. The transition probability

$$\Pr\Big[N(t_0 + t) = n \,|\, N(t_0) = m\Big]$$

is the conditional probability that process $N(t)$ has moved from state m to state n during the time interval $[t_0, t_0 + t]$. The analysis of a stochastic process becomes greatly simplified when the above transition probability depends on: i) the starting state m but not on the states before time t_0 and ii) the interval-length t but not on the start time t_0. Property (i) is the well-known *Markov property* and the process with this property is said to be a *Markov process*. The process holding property (ii) is said to be a *homogeneous process*. If the molecules are well mixed and are available everywhere for a reaction (space can be ignored), then the copy number $N(t)$ can be approximately formulated as a homogeneous Markov process in continuous time. In this text, all Markov processes will be assumed to be homogeneous unless stated otherwise. Now we use a simple notation for the above transition probability

$$P(n|m,t) = \Pr\Big[N(t_0 + t) = n \,|\, N(t_0) = m\Big] = \Pr\Big[N(t) = n \,|\, N(0) = m\Big]. \tag{2.17}$$

It should be remembered that t in the above equation is the length of the time interval. The initial condition is usually fixed and the state probability can be written as a transition probability

$$P(n,t) = P(n|n^0,t) = \Pr\Big[N(t) = n \,|\, N(0) = n^0\Big].$$

The Markov property has two important consequences, explained in the following two sections.

2.3.1 Chapman-Kolmogorov equation

The Markov property places a consistency condition on the transition probabilities. To see that, decompose the transition probability

$$\begin{aligned}
&\Pr\Big[X(t+w) = n \,|\, X(0) = m\Big] \\
&\qquad = \sum_{n'} \Pr\Big[X(t+w) = n \,|\, X(t) = n' \cap X(0) = m\Big] \Pr\Big[X(t) = n' \,|\, X(0) = m\Big] \\
&\qquad = \sum_{n'} \Pr\Big[X(t+w) = n \,|\, X(t) = n'\Big] \Pr\Big[X(t) = n' \,|\, X(0) = m\Big]
\end{aligned}$$

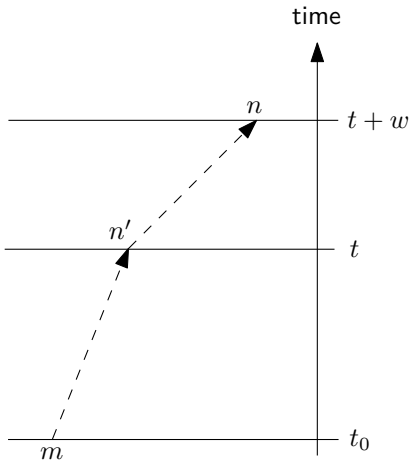

Figure 2.1
Graphical interpretation of the Chapman-Kolmogorov equation. The probability of a transition $m \to n$ can be obtained by summing up the probabilities of all transitions $m \to n' \to n$, via intermediate states n'. Drawing adopted from [Gil92].

where the Markov property allows a simplification of the 2nd line leading to the 3rd line. In the compact notation for transition probabilities, the above consistency condition takes the form

$$P(n|m, t+w) = \sum_{n'} P(n|n', w)P(n'|m, t), \tag{2.18}$$

which is known as the *Chapman-Kolmogorov equation* (CKE) for continuous-time Markov processes. This equation expresses the probability of a transition $(m \to n)$ as the summation of probabilities of all transitions $(m \to n' \to n)$ via the intermediate states n'. Figure 2.1 illustrates the idea conveyed by the CKE. It is important to clarify that the CKE is only a consistency condition imposed on every stochastic process by Markov property and cannot characterise a particular process. We need dependence relations between random variables of the process to characterise it. Typically that is achieved by investigating the local behaviour of transition probabilities in a short time interval. Replacing the length w of the time interval of the transition probabilities in (2.18) by Δt and fixing the initial condition, the CKE (2.18) reduces to

$$P(n, t+\Delta t) = \sum_{n'} P(n|n', \Delta t)P(n', t), \tag{2.19}$$

where the transition probabilities away from the fixed initial state have been replaced by the state probabilities. Later we will see that the short-time transition probabilities $P(n|n', \Delta t)$ can be expressed in terms of parameters of the particular process under consideration. This will open the door for an analytical characterisation of a particular Markov process.

2.3.2 Memoryless property

Suppose the Markov process $N(t)$ is in state n at time t_0 and let $T_j(n)$ denote the waiting time in state n until the occurrence of a reaction R_j takes the process to state $n + S_{\cdot j}$. If the reaction has not occurred during $[t_0, t_0 + w]$, we can write $T_j(n) > w$. This knowledge, however, does not change the uncertainty in time until the next reaction. In other words, the process is *memoryless* and its subsequent behaviour is independent of w.

Mathematically, this memoryless property is written as

$$\Pr\Big[T_j(n) > w + t \,|\, T_j(n) > w\Big] = \Pr\Big[T_j(n) > t\Big]\,. \tag{2.20}$$

and holds true only for the exponential distribution, as seen in the following.

Exponential distribution: Consider a non-negative continuous random variable W satisfying the memoryless property. Then the complementary cumulative distribution function (CCDF)

$$G(t) = \Pr\Big[W > t\Big]$$

satisfies

$$\begin{aligned} G(w+t) &= \Pr\Big[W > w + t\Big] \\ &= \Pr\Big[W > w + t \cap W > w\Big] \\ &= \Pr\Big[W > w\Big] \Pr\Big[W > w + t \,|\, W > w\Big] \\ &= \Pr\left[W > w\right] \Pr\left[W > t\right] \\ &= G(w)G(t)\,. \end{aligned}$$

Putting $w = 0$ gives $G(t) = G(0)G(t)$ which has the nontrivial implication $G(0) = 1$. The derivatives of $G(w+t)$ with respect to w and t are equal,

$$\frac{\mathrm{d}G(w)}{\mathrm{d}w}G(t) = G(w)\frac{\mathrm{d}G(t)}{\mathrm{d}t}$$

which can be rearranged to

$$\frac{1}{G(w)}\frac{\mathrm{d}G(w)}{\mathrm{d}w} = \frac{1}{G(t)}\frac{\mathrm{d}G(t)}{\mathrm{d}t} = -\lambda,$$

where λ is a constant. The nontrivial solution $G(t) = \exp(-\lambda t)$ is bounded for $\lambda > 0$ and corresponds to the exponential random variable. Hence it follows that any non-negative continuous random variable W satisfying the memoryless property (2.20) is exponentially distributed.

The memoryless property, and hence the fact that the times between reactions are exponentially distributed, opens the door for stochastic simulations of biochemical reaction networks. That will be our focus in the following section.

2.4 Propensity as the transition rate

It follows from the previous section that the time $T_j(n)$ until the occurrence of reaction R_j has an exponential distribution with a parameter, say $a_j(n)$. We can thus write

$$\Pr\Big[T_j(n) > t\Big] = \exp\Big(-a_j(n)t\Big), \tag{2.21}$$

for the probability that R_j will not occur in the next time interval of length t. Using a Taylor series expansion, for arbitrarily short interval of length Δt, the above probability can be written as

$$\Pr\Big[T_j(n) > \Delta t\Big] = \exp\Big(-a_j(n)\Delta t\Big) = 1 - a_j(n)\Delta t + o(\Delta t)\,. \tag{2.22}$$

The probability of occurrence of R_j during the same short interval is complimentary to the above:

$$\Pr\Big[T_j(n) \le \Delta t\Big] = a_j(n)\Delta t + o(\Delta t)\,. \tag{2.23}$$

The parameter $a_j(n)$, which gives the probability per unit time of the occurrence of R_j in state n, is referred to as the *reaction propensity*.

In a vanishingly short interval, it is highly improbable that a particular reaction will occur more than once. To see that, the probability of two occurrences of R_j during a time interval $[t, t+\Delta t]$ is the joint probability of its first occurrence during $[t, t+\alpha\Delta t]$ and a second occurrence during $(t+\alpha\Delta t, t+\Delta t]$:

$$\begin{aligned}\Pr\Big[T_j(n) \le \alpha\Delta t\Big]\, &\Pr\Big[T_j\left(n + S_{\bullet j}\right) \le (1-\alpha)\Delta t\Big] \\ &= \Big(a_j(n)\alpha\Delta t + o(\Delta t)\Big)\Big(a_j\left(n + S_{\bullet j}\right)(1-\alpha)\Delta t + o(\Delta t)\Big) = o(\Delta t),\end{aligned}$$

where $0 < \alpha < 1$. Therefore, the probability in (2.23) is equivalent to the probability, in state n, of *one* occurrence (i.e. a unit increment in the DA) of R_j during $[t, t+\Delta t]$:

$$\Pr\Big[Z_j(t+\Delta t) - Z_j(t) = 1 \,|\, N(t) = n\Big] = a_j(n)\Delta t + o(\Delta t)\,.$$

The probability distribution, in state n, of the random progress $Z_j(t+\Delta t) - Z_j(t)$ of the jth reaction during $[t, t+\Delta t)$ is

$$\Pr\Big[Z_j(t+\Delta t) - Z_j(t) = z_j \,|\, N(t) = n\Big] = o(\Delta t) + \begin{cases} a_j(n)\Delta t & \text{if } z_j = 1 \\ 1 - a_j(n)\Delta t & \text{if } z_j = 0 \\ 0 & \text{if } z_j > 1 \end{cases} \tag{2.24}$$

The expected value of this short-time DA increment is

$$\begin{aligned}&\langle Z_j(t+\Delta t) - Z_j(t) \,|\, N(t) = n\rangle \\ &\qquad = \sum_{j=0}^{r} z_j \Pr\left[Z_j(t+\Delta t) - Z_j(t) = z_j \,|\, N(t) = n\right] == \overbrace{a_j(n)\Delta t}^{z_j=1} + \overbrace{o(\Delta t)}^{z_j>1}\end{aligned}$$

which is conditioned on $N(t) = n$. The unconditional expectation of the DA increment can be obtained by summing the probabilities $P(n,t)$ weighted by the above conditional expectation over all possible states n:

$$\begin{aligned}\Big\langle Z_j(t+\Delta t) - Z_j(t)\Big\rangle &= \sum_n \Big\langle Z_j(t+\Delta t) - Z_j(t) \,|\, N(t) = n\Big\rangle P(n,t) \\ &= \sum_n a_j(n) P(n,t)\Delta t + o(\Delta t) \\ &= \langle a_j(N(t))\rangle\, \Delta t + o(\Delta t)\end{aligned}$$

which for vanishingly small Δt leads to the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Z_j(t)\rangle = \langle a_j(N(t))\rangle\,, \tag{2.25}$$

Thus the mean propensity of a particular reaction can be interpreted as the *average number of occurrences per unit time* of that reaction.

The state transition associated with R_j will be written as

$$n \xrightarrow{a_j(n)} n + S_{\bullet j}\,.$$

The dependence relation of the propensity on the state n is determined by the system being modelled and reflects the assumptions made about the system. If R_j is an elementary reaction in a well-mixed system, it is reasonable to assume that each possible combination of the R_j reactant molecules has the same probability per unit time, c_j, to react. In other words $c_j dt$ gives the probability that a particular combination of R_j reactant molecules will react in a short time interval $(t, t+dt]$. In the literature, c_j is referred to as *stochastic reaction rate constant*. If there are $h_j(n)$ different possible combinations of R_j reactant molecules in state n, then the propensity $a_j(n)$ can be written as

$$a_j(n) = c_j h_j(n)\,. \tag{2.26}$$

The form of $h_j(n)$ depends on the order of the reaction R_j. For an elementary reaction channel of the general form (2.1) we can write the combinatorial function

$$h_j(n) = \prod_{i=1}^{s} \binom{n_i}{S_{ij}}\,. \tag{2.27}$$

However, it is highly unlikely that a reaction of order higher than two will result from all its reactants coming together and reacting in one step, for example by collision. A more realistic model will decompose the high order reaction into two or more one step reactions. For non-elementary reactions, the propensity can be computed from the reaction rate by using (2.28). For elementary reactions, the stochastic rate constant c is closely related to the deterministic rate constant, as shown below.

Deterministic and stochastic reaction rates: Using the interpretation of propensity as the mean DA per unit time from (2.25), the propensity divided by the system size is analogous to the reaction rate v_j defined earlier in the deterministic framework. Hence, in the stochastic framework, the *stochastic reaction rate* can be defined as

$$\tilde{a}(x) = \frac{a(n)}{\Omega}\,. \tag{2.28}$$

which is analogous to the deterministic reaction rate $v(x)$. The stochastic rate of a given elementary reaction can be computed from (2.28), (2.26) and (2.27) whereas (2.12) can be used for the deterministic reaction rate. The two kinds of reaction rates are given for a few example elementary reactions in Table 2.1. The condition under which the two reaction rates are equal is shown in the corresponding entry of the last column. This also provides the relationship between the stochastic rate constant c_j and the deterministic rate constant k_j. That relationship can be generalised in the following way.

Table 2.1 Examples of elementary reactions and their reaction rates.

R_j	$a_j(n)$	$\tilde{a}(x)$	$v(x)$	$\tilde{a}(x) = v(x)$ if
$\varnothing \xrightarrow{k_j} X_1$	c_j	$\frac{c_j}{\Omega}$	k_j	$c_j = \Omega k_j$
$X_1 \xrightarrow{k_j} ?$	$c_j n_1$	$\frac{c_j n_1}{\Omega}$	$k_j x_1$	$c_j = k_j$
$X_1 + X_2 \xrightarrow{k_j} ?$	$c_j n_1 n_2$	$\frac{c_j n_1 n_2}{\Omega}$	$k_j x_1 x_2$	$c_j = \frac{k_j}{\Omega}$
$2X_1 \xrightarrow{k_j} ?$	$c_j \frac{(n_1-1)n_1}{2}$	$c_j \frac{(n_1-1)n_1}{2\Omega}$	$k_j x_1^2$	$c_j = \frac{2k_j}{\Omega},\ \Omega \gg 1$
$X_1 + X_2 + X_3 \xrightarrow{k_j} ?$	$c_j n_1 n_2 n_3$	$c_j \frac{n_1 n_2 n_3}{\Omega}$	$k_j x_1 x_2 x_3$	$c_j = \frac{k_j}{\Omega^2}$
$X_1 + 2X_2 \xrightarrow{k_j} ?$	$c_j n_1 \frac{(n_2-1)n_2}{2}$	$c_j \frac{(n_2-1)n_2 n_1}{2\Omega}$	$k_j x_1 x_2^2$	$c_j = \frac{2k_j}{\Omega^2},\ \Omega \gg 1$

Relationship between the deterministic and stochastic rate constants: Let us find the conditions under which the deterministic and stochastic reaction rates of a general elementary reaction are approximately the same. From (2.12), (2.28), (2.26) and (2.27) we can propose:

$$k_j \prod_{i=1}^{s} x_i^{\underline{S}_{ij}} = v_j(x) \approx \tilde{a}_j(x) = \frac{a_j(n)}{\Omega} = \frac{c_j}{\Omega} \prod_{i=1}^{s} \binom{n_i}{\underline{S}_{ij}}.$$

The left-most expression is valid only in the deterministic framework which requires large system size, $\Omega \gg 1$. To the extent that this assumption is valid, the combinatorial function can be approximated as

$$\begin{aligned}
\binom{n_i}{\underline{S}_{ij}} &= \frac{(n_i - \underline{S}_{ij} + 1) \cdots (n_i - 1)\, n_i}{\underline{S}_{ij}!} \\
&= \left(\frac{\Omega^{\underline{S}_{ij}}}{\underline{S}_{ij}!}\right) \left(x_i - \frac{\underline{S}_{ij} - 1}{\Omega}\right) \cdots \left(x_i - \frac{1}{\Omega}\right) x_i \\
&\approx \left(\frac{\Omega^{\underline{S}_{ij}}}{\underline{S}_{ij}!}\right) x_i^{\underline{S}_{ij}} \quad \text{for} \quad \Omega \gg 1
\end{aligned}$$

Inserted into the previous equation leads to the stochastic rate constant

$$c_j = \frac{k_j}{\Omega^{K_j - 1}} \prod_{i=1}^{s} (\underline{S}_{ij}!) \tag{2.29}$$

where $K_j = \sum_{i=1}^{s} \underline{S}_{ij}$ is the number of R_j reactant molecules required to collide and possibly result in a single occurrence of the reaction. The above derivation is a refinement of our earlier attempt in [WUKC04].

Relation between the deterministic and stochastic reaction rates: We saw that in general, the stochastic and deterministic reaction rates are not equal. Since the two are

equal for infinitely large Ω, the difference between the two is of the order of Ω^{-1}, namely [Elf04]

$$\tilde{a}_j(x) = v_j(x) + \mathcal{O}\left(\Omega^{-1}\right) . \tag{2.30}$$

Isomerisation reaction revisited: In the isomerisation reaction (2.5), the copy number $N(t)$ of the protein in the inactive form is a simple birth death process. Each copy of the (inactive) protein X_1 is activated at a rate k_1. Similarly, each copy of the (active) protein X_2 is deactivated at a rate k_2. With $0 < n < n^{\text{tot}}$ proteins in the inactive form X_2, the over all activation (death) rate is the sum $a_1(n) = k_1 n$ and the overall deactivation (birth) rate is the sum $a_2(n) = (n^{\text{tot}} - n)k_2$. The state transitions in state n are listed here (on the right) together with the corresponding reactions (on the left)

$$\left.\begin{array}{l|l} \mathrm{X}_1 \xrightarrow{k_1} \mathrm{X}_2 & n \xrightarrow{k_1 n} n-1 \\ \mathrm{X}_2 \xrightarrow{k_2} \mathrm{X}_1 & n \xrightarrow{(n^{\text{tot}}-n)k_2} n+1 \end{array}\right\} \tag{2.31}$$

Lotka-Volterra model revisited: The predator-prey model (2.6) can be formulated as a 2-component 3-reaction network. Let $N_1(t)$ denote the population of the prey, and $N_2(t)$ that of the predator. Then the Markov process $N(t) = (N_1(t), N_2(t))^T$ has states $n = (n_1, n_2)^T$. State transitions in state n are listed here (on the right) together with the corresponding reactions (on the left):

$$\left.\begin{array}{l|l} \mathrm{X}_1 + \mathrm{A} \xrightarrow{k_1} 2\mathrm{X}_1 & n \xrightarrow{k_1 n_{\mathrm{A}} n_1} (n_1+1, n_2)^T \\ \mathrm{X}_1 + \mathrm{X}_2 \xrightarrow{k_2} 2\mathrm{X}_2 & n \xrightarrow{k_2 n_1 n_2} (n_1-1, n_2+1)^T \\ \mathrm{X}_2 \xrightarrow{k_3} \varnothing & n \xrightarrow{k_3 n_2} (n_1, n_2-1)^T \end{array}\right\} \tag{2.32}$$

Enzyme kinetic model revisited: The enzyme kinetic model (2.8) is a 4-component 3-reaction network. Let $N_{\mathrm{E}}(t)$ denote the copy number of the enzyme, $N_{\mathrm{S}}(t)$ that of the substrate, $N_{\mathrm{ES}}(t)$ that of the complex and $N_{\mathrm{P}}(t)$ that of the product. The full state has to respect two conservation laws

$$N_{\mathrm{E}}(t) + N_{\mathrm{ES}}(t) = n_{\mathrm{E}}^{\text{tot}}, \quad \text{and} \quad N_{\mathrm{S}}(t) + N_{\mathrm{ES}}(t) + N_{\mathrm{P}}(t) = n_{\mathrm{S}}^{\text{tot}}$$

where $n_{\mathrm{E}}^{\text{tot}}$ and $n_{\mathrm{S}}^{\text{tot}}$ are the total copy numbers of the enzyme and substrate, respectively. The Markov process

$$N(t) = \left(N_{\mathrm{S}}(t), N_{\mathrm{ES}}(t)\right)^T$$

having states $n = (n_{\mathrm{S}}, n_{\mathrm{ES}})^T$ is sufficient to describe the system, because the remaining two variables can be determined from the conservation relations above. The transitions in state n are listed here (on the right) together with the corresponding reactions (on the

left):

$$\left.\begin{array}{l|l} \mathrm{E}+\mathrm{S} \xrightarrow{k_1} \mathrm{ES} & n \xrightarrow{\left(\frac{k_1}{\Omega}\right)\left(n_{\mathrm{S}}^{\mathrm{tot}}-n_{\mathrm{ES}}\right)n_{\mathrm{S}}} \left(n_{\mathrm{S}}-1, n_{\mathrm{ES}}+1\right)^T \\ \mathrm{ES} \xrightarrow{k_2} \mathrm{E}+\mathrm{S} & n \xrightarrow{k_2 n_{\mathrm{ES}}} \left(n_{\mathrm{S}}+1, n_{\mathrm{ES}}-1\right)^T \\ \mathrm{ES} \xrightarrow{k_3} \mathrm{E}+\mathrm{P} & n \xrightarrow{k_3 n_{\mathrm{ES}}} \left(n_{\mathrm{S}}, n_{\mathrm{ES}}-1\right)^T \end{array}\right\} \tag{2.33}$$

where $\Omega = N_{\mathrm{A}} V$.

Schlögl model revisited: For the Schlögl reaction scheme (2.8), write x_{A} and x_{B} denote the constant respective concentrations of chemicals A and B, and $N(t)$ for the time-dependent copy number of chemical X. State transitions in state n are listed here (on the right) together with the corresponding reactions (on the left):

$$\left.\begin{array}{l|l} \left.\begin{array}{r} \mathrm{A}+2\mathrm{X} \xrightarrow{k_1} 3\mathrm{X} \\ \mathrm{B} \xrightarrow{k_3} \mathrm{X} \end{array}\right\} & n \xrightarrow{\left(\frac{k_1 x_{\mathrm{A}}}{\Omega}\right)(n-1)n+k_3 x_{\mathrm{B}}\Omega} n+1 \\ \left.\begin{array}{r} 3\mathrm{X} \xrightarrow{k_2} \mathrm{A}+2\mathrm{X} \\ \mathrm{X} \xrightarrow{k_4} \mathrm{B} \end{array}\right\} & n \xrightarrow{\left(\frac{k_2}{\Omega^2}\right)(n-2)(n-1)n+k_4 n} n-1 \end{array}\right\} \tag{2.34}$$

where $\Omega = N_{\mathrm{A}} V$.

2.5 Stochastic simulation

Time until the next reaction: The probability in state n that no reaction has occurred in an interval of length t follows from (2.21) and independence of reaction channels:

$$\Pr\left[\bigcap_j T_j(n) > t\right] = \prod_j \exp\left(-a_j(n)t\right) = \exp\left(-t\sum_j a_j(n)\right).$$

Hence the time $T_0(n)$ until next reaction taking the process away from state n is exponential with rate parameter

$$a_0(n) = \sum_j a_j(n),$$

known as the *exit rate* of state n.

Index of the next reaction channel: If it is known that a reaction has occurred in state n, the (conditional) probability that it was an R_j reaction is determined as

$$\begin{aligned}\lim_{\Delta t\to 0}\Pr\Big[T_j(n)\le\Delta t\,|\,T_0(n)\le\Delta t\Big] &= \lim_{\Delta t\to 0}\frac{\Pr\Big[T_j(n)\le\Delta t\Big]}{\Pr\Big[T_0(n)\le\Delta t\Big]}\\ &= \lim_{\Delta t\to 0}\frac{a_j(n)\Delta t+o(\Delta t)}{a_0(n)\Delta t+o(\Delta t)}\\ &= \frac{a_j(n)}{a_0(n)}\,.\end{aligned}$$

Thus the index $J(n)$ of the next reaction known to have occurred in state n is a discrete random variable taking values j with probability

$$\Pr\Big[J(n)=j\Big]=\frac{a_j(n)}{a_0(n)}\,. \tag{2.35}$$

This result, together with the exponentially of T_0, allows a simple procedure to simulate the Markov process: 1) Pick a sample τ from the exponential distribution with rate $a_0(n)$ to realise the time until the next reaction will occur, and 2) pick a sample j from the discrete distribution with probabilities (2.35) to realise the reaction channel.

Random number generation: Here is a brief review of one way to compute a sample y of a random variable Y. The sample y is a random number picked from the probability distribution of Y. We start with the cumulative distribution function (CDF) defined as

$$F(y)=\Pr\Big[Y\le y\Big]\,.$$

The CDF of the transformed random variable $U=F(Y)$ can be worked out

$$\Pr\Big[U\le u\Big]=\Pr\Big[Y\le F^{-1}(u)\Big]=F\left(F^{-1}(u)\right)=u,$$

which is the CDF of the uniform random variable in the interval $[0,1]$. Thus U is uniform in the unit interval. The same holds true for the CCDF, defined as

$$G(y)=\Pr\Big[Y>y\Big]=1-F(y)\,.$$

If u is a uniform random number picked from the unit interval, then $y=F^{-1}(u)$ is a sample of Y. Another sample of Y is $G^{-1}(u)$. Computation of the inverse $F^{-1}(u)$ is straightforward when a closed form expression of F is available. When the distribution has no unique inverse, the inverse is defined by

$$y=F^{-1}(u)=\min_w\left\{w:F(w)\ge u\right\}\,.$$

Simulating the time until the next reaction: The time $T_0(n)$ until the next reaction in state n is an exponential random variable with right-tailed distribution function

$$G(t) = \Pr\Big[T_0(n) > t\Big] = \exp\Big(-a_0(n)t\Big)\,.$$

If u_1 is a uniform random number picked from $[0, 1]$, then

$$\tau = G^{-1}(u_1) = -\frac{\log u_1}{a_0(n)}. \tag{2.36}$$

is a sample of the time until the next reaction.

Simulating the Index of the next reaction channel: The index $J(n)$ of the next reaction known to have occurred in state n is a discrete random variable with a probability mass function (2.35) and a cumulative distribution function

$$F(j) = \Pr\Big[J(n) \le j\Big] = \sum_{l=1}^{j} \frac{a_l(n)}{a_0(n)}\,.$$

If u_2 is a uniform random number picked from $[0, 1]$ then

$$j = F^{-1}(u_2) = \min_w \{w : F(w) \ge u_2\}$$

is a sample of the index $J(n)$. For the range of values taken by J, the above condition is equivalent to

$$F(j-1) < u_2 \le F(j)\,.$$

Multiplying both sides by $a_0(n)$ and plugging values for $F(j)$ gives the following criteria

$$\sum_{l=1}^{j-1} a_l(n) < u_2 a_0(n) \le \sum_{l=1}^{j} a_l(n)\,. \tag{2.37}$$

for j to be a sample of the index $J(n)$ of the next reaction known to have occurred in state n.

Gillespie algorithm [Gil77]: The above two results (2.36) and (2.37) are at the core of the stochastic simulation algorithm also known as the "Gillespie algorithm". The steps involved are listed in Algorithm 2.1. Over time, many improvements to the original SSA have been made for efficient computation. See [TSB04, Pah08] for extensive reviews.

2.6 Chemical master equation

How does the state probability $P(n, t)$ change with time? To answer this, we need to find an expression for $P(n, t + \Delta t)$, the probability to be in state n after a short time-interval of length Δt. How can the system fall in state n at time $t + \Delta t$? One possibility is that the system was in state n at time t and no reaction occurred during the interval. Otherwise

Algorithm 2.1 Gillespie's stochastic simulation algorithm (direct method)

1. Initialise the system at $t = 0$ with initial numbers of molecules for each species, $n_1, \ldots, n_s$.
2. For each $j = 1, \ldots, r$, calculate $a_j(n)$ based on the current state n.
3. Calculate the exit-rate $a_0(n) = \sum_{j=1}^{r} a_j(n)$. Terminate if $a_0(n) = 0$.
4. Compute a sample τ of the time until the next reaction using (2.36).
5. Update the time $t = t + \tau$.
6. Compute a sample j of the reaction index using (2.37).
7. Update the state n according to R_j. That is put $n = n + S_{\cdot j}$, where $S_{\cdot j}$ denotes jth column of the stoichiometric matrix S.
8. If $t < t_{\max}$, return to Step 2.

the state n was reached after the occurrence of one of r possible reactions. Mathematically we can write

$$P(n|n', \Delta t) = o(\Delta t) + \begin{cases} 1 - a_0(n)\Delta t & \text{if } n' = n \\ a_1\left(n - S_{\cdot 1}\right)\Delta t & \text{if } n' = n - S_{\cdot 1} \\ \vdots & \\ a_r\left(n - S_{\cdot r}\right)\Delta t & \text{if } n' = n - S_{\cdot r} \\ 0 & \text{elsewhere.} \end{cases}$$

The term $o(\Delta t)$ represents the probability of arriving in state n by the occurrence of more than one reaction during the interval. Recall that $a_0(n) = \sum_j a_j(n)$ is the exit-rate from state n. Substituting the above expressions into (2.19) gives

$$P(n, t + \Delta t) = P(n, t)\left(1 - \sum_{j=1}^{r} a_j(n)\Delta t\right) + \sum_{j=1}^{r} P\left(n - S_{\cdot j}, t\right) a_j\left(n - S_{\cdot j}\right)\Delta t + o(\Delta t),$$

which for vanishingly short Δt can be re-arranged to what is known as the *chemical master equation* (CME):

$$\frac{\partial}{\partial t}P(n, t) = \sum_{j=1}^{r}\Big[a_j\left(n - S_{\cdot j}\right)P\left(n - S_{\cdot j}, t\right) - a_j(n)P(n, t)\Big]. \tag{2.38}$$

We will switch between the two alternative notations $\frac{\mathrm{d}}{\mathrm{d}t}\phi(t)$ and $\frac{\mathrm{d}\phi}{\mathrm{d}t}$ for any scalar quantity $\phi(t)$. We will prefer the later when dependence on time variable is implicitly clear.

Using a negative-shift operator $\mathbb{E}_j$ for each reaction channel defined by its effect

$$\mathbb{E}_j\, f(n) = f\left(n + S_{\cdot j}\right)$$

on an arbitrary scalar function $f(n)$ of s-vector n, the CME can be written in the alternative form

$$\frac{\partial}{\partial t}P(n,t) = \sum_{j=1}^{r}\left(\mathbb{E}_j^{-1} - 1\right) a_j(n)P(n,t)\,. \tag{2.39}$$

Since there is one equation for each state n and there is potentially a large number of possible states, it is impractical to solve the CME.

Isomerisation revisited:I Following from the state transitions (2.31) for the isomerisation reaction scheme (2.5), the channel-wise summands in the CME are:

j	$\left(\mathbb{E}_j^{-1} - 1\right) a_j(n)P(n,t)$
1	$k_1\Big[(n+1)P(n+1,t) - nP(n,t)\Big]$
2	$k_2\Big[\left(n^{\text{tot}} - n + 1\right)P(n-1,t) - \left(n^{\text{tot}} - n\right)P(n,t)\Big]$

Enzyme kinetic reaction revisited: Following the state transitions (2.33) for the enzyme kinetic reaction scheme (2.7), the channel-wise summands in the CME are:

j	$\left(\mathbb{E}_j^{-1} - 1\right) a_j\left(n_{\text{S}}, n_{\text{ES}}\right) P\left(n_{\text{S}}, n_{\text{ES}}, t\right)$
1	$\left(\frac{k_1}{\Omega}\right)\left[\left(n_{\text{S}}^{\text{tot}} - n_{\text{ES}} + 1\right)\left(n_{\text{S}} + 1\right)P\left(n_{\text{S}} + 1, n_{\text{ES}} - 1, t\right) - \left(n_{\text{S}}^{\text{tot}} - n_{\text{ES}}\right)n_{\text{S}}P\left(n_{\text{S}}, n_{\text{ES}}, t\right)\right]$
2	$k_2\Big[\left(n_{\text{ES}} + 1\right)P\left(n_{\text{S}} - 1, n_{\text{ES}} + 1, t\right) - n_{ES}P\left(n_{\text{S}}, n_{\text{ES}}, t\right)\Big]$
3	$k_3\Big[\left(n_{\text{ES}} + 1\right)P\left(n_{\text{S}}, n_{\text{ES}} + 1, t\right) - n_{ES}P\left(n_{\text{S}}, n_{\text{ES}}, t\right)\Big]$

Lotka-Volterra model revisited: Following the state transitions (2.32) for the predator-prey reaction scheme (2.6), the channel-wise summands in the CME are:

j	$\left(\mathbb{E}_j^{-1} - 1\right) a_j\left(n_1, n_2\right) P\left(n_1, n_2, t\right)$
1	$k_1\Big[\left(n_1 - 1\right)P\left(n_1 - 1, n_2, t\right) - n_1P\left(n_1, n_2, t\right)\Big]$
2	$k_2\Big[\left(n_1 + 1\right)\left(n_2 - 1\right)P\left(n_1 + 1, n_2 - 1, t\right) - n_1n_2P\left(n_1, n_2, t\right)\Big]$
3	$k_3\Big[\left(n_2 + 1\right)P\left(n_1, n_2 + 1, t\right) - n_2P\left(n_1, n_2, t\right)\Big]$

Schlögl model revisited: Following the state transitions (2.34) for the Schlögl reaction scheme (2.8), the channel-wise summands in the CME are:

j	$\left(\mathbb{E}_j^{-1}-1\right) a_j(n)P(n,t)$
1	$\left[\left(\frac{k_1 x_{\mathrm{A}}}{\Omega}\right)(n-2)(n-1)+k_3 x_{\mathrm{B}}\Omega\right]P(n-1,t)-\left[\left(\frac{k_1 x_{\mathrm{A}}}{\Omega}\right)(n-1)n+k_3 x_{\mathrm{B}}\Omega\right]P(n,t)$
2	$\left(\frac{k_2}{\Omega^2}\right)\Big[(n+1)P(n+1,t)-(n-2)P(n,t)\Big](n-1)n+k_4\Big[(n+1)P(n+1,t)-nP(n,t)\Big]$

While the stochastic simulation algorithm and extensions provide a way to generate sample paths of copy numbers for a biochemical system, the need for repeating many simulation runs to get an idea of the probability distribution in terms of its moments (mean and (co)variance) become increasing time consuming and even impractical for larger systems. Therefore attempts have been made towards approximations of the CME [Gil96, HJ04, PMK06, MK06], including the following.

2.6.1 Kramers-Moyal expansion and the Fokker-Planck equation

Suppose the propensity $a_j(n)$ is a smooth function and one is interested in solutions $P^c(n,t)$ that can be represented by a smooth function. It is then reasonable to approximate the problem by means of a description in which n is treated as a continuous variable. The operator $\mathbb{E}_j^{-1}$, acting only on smooth functions, may be replaced with a Taylor expansion,

$$\begin{aligned}\mathbb{E}_j^{-1} &= \sum_{m=0}^{\infty}\frac{1}{m!}\left(-\sum_i S_{ij}\frac{\partial}{\partial n_i}\right)^m \\ &= 1-\sum_i S_{ij}\frac{\partial}{\partial n_i}+\frac{1}{2}\sum_{i,k}S_{ij}S_{kj}\frac{\partial^2}{\partial n_i\partial n_k}+\cdots .\end{aligned} \tag{2.40}$$

Inserting values into the master equation yields the Kramers-Moyal expansion,

$$\frac{\partial}{\partial t}P^c(n,t)=\sum_{j=1}^{r}\sum_{m=1}^{\infty}\frac{1}{m!}\left(-\sum_i S_{ij}\frac{\partial}{\partial n_i}\right)^m a_j(n)P^c(n,t)\,.$$

Ignoring all derivatives beyond the second yields the Fokker-Planck equation (FPE),

$$\frac{\partial}{\partial t}P^c(n,t)=\sum_{j=1}^{r}\left(-\sum_i S_{ij}\frac{\partial}{\partial n_i}+\frac{1}{2}\sum_{i,k}S_{ij}S_{kj}\frac{\partial^2}{\partial n_i\partial n_k}\right)a_j(n)P^c(n,t)\,. \tag{2.41}$$

2.6.2 System-size expansion

The cutting of higher moments in the Kramers-Moyal expansion to get the Fokker-Planck approximation requires that the fluctuations, as measured by the standard deviations σ, are small. However changes in copy numbers by chemical reactions are whole numbers and there is no way to qualify small fluctuations in this way. This is true for any approximation method that requires small fluctuations. Therefore one needs a systematic approximation method in the form of an expansion in powers of a small parameter. Only in that case does one have an objective measure for the size of the several terms. The expansion parameter must appear in the master equation and must govern the size of the fluctuations. The system size parameter Ω is a potential choice. Suppose the propensity $a_j(n)$ is a smooth function and one is interested in solutions $P^c(n,t)$ that can be represented by a smooth function. It is then reasonable to approximate the problem by means of a description in which n is treated as a continuous variable. We can anticipate the way in which the solution $P^c(n,t)$ will depend on the system size Ω. The initial condition is

$$P^c(n,t) = \delta(n - n^0)\,.$$

The initial copy number $X(0) = n^0 = \Omega x(0)$ is of order Ω. The Dirac's delta function $\delta(n - n^0)$ is defined to be zero everywhere except at $n = n_0$ where it integrates to unity. One expects that at later times $P^c(n,t)$ is a sharp peak at some position of order Ω while its width will be of the order $\sqrt{\Omega}$. In order words it is assumed that the continuous approximation $N^c(t)$ of the process $N(t)$ fluctuates around a macroscopic trajectory of the order Ω with a fluctuation of order $\Omega^{1/2}$. To express this formally we set

$$N^c(t) = \Omega\phi(t) + \Omega^{1/2}\Xi(t), \tag{2.42}$$

where $\phi(t)$ is equal to the macroscopic concentration $x = n/\Omega$ for an infinitely large system size Ω and $\Xi(t)$ models the fluctuation of $N^c(t)$ around $\phi(t)$. A realisation n of $N^c(t)$ is related to a realisation ξ of $\Xi(t)$ by the same relation above:

$$n = \Omega\phi(t) + \Omega^{1/2}\xi\,.$$

The probability distribution $P^c(n,t)$ of $N^c(t)$ transforms into a probability distribution $\Pi(\xi,t)$ of $\Xi(t)$ according to

$$P^c(n,t) = P^c\left(\Omega\phi(t) + \Omega^{1/2}\xi, t\right) = \Pi(\xi,t)\,. \tag{2.43}$$

The time derivative in the master equation is taken with constant n, that is where,

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \Omega\frac{\mathrm{d}\phi}{\mathrm{d}t} + \Omega^{1/2}\frac{\mathrm{d}\xi}{\mathrm{d}t} = 0 \quad \Longrightarrow \quad \frac{\mathrm{d}\xi}{\mathrm{d}t} = -\Omega^{1/2}\frac{\mathrm{d}\phi}{\mathrm{d}t}\,.$$

This result can be used in the differentiation of the probability distributions with respect to time to give

$$\frac{\partial}{\partial t}P^c(n,t) = \frac{\partial}{\partial t}\Pi(\xi,t) + \sum_{i=1}^{s}\frac{\mathrm{d}\xi_i}{\mathrm{d}t}\frac{\partial}{\partial\xi_i}\Pi(\xi,t) = \frac{\partial}{\partial t}\Pi(\xi,t) - \Omega^{1/2}\sum_{i=1}^{s}\frac{\mathrm{d}\phi_i}{\mathrm{d}t}\frac{\partial}{\partial\xi_i}\Pi(\xi,t)\,. \tag{2.44}$$

Before we can compare this equation with the CME (2.39), we need to express the propensity function $a_j(n)$ in terms of the fluctuation ξ and translate the operator $\mathbb{E}_j$ so

that it can be applied to functions of ξ. The propensity is related to the deterministic reaction rate $v_j(x)$ through (2.28) and (2.30):

$$a_j(n) = \Omega\tilde{a}_j\left(\frac{n}{\Omega}\right) = \Omega\tilde{a}_j\left(\phi + \Omega^{-1/2}\xi\right) = \Omega\Big[v_j\left(\phi + \Omega^{-1/2}\xi\right) + \mathcal{O}\left(\Omega^{-1}\right)\Big].$$

The operator $\mathbb{E}_j^{-1}$ which changes n to $n - S_{\cdot j}$, effectively changing the fluctuation ξ to $\xi - \Omega^{-1/2}S_{\cdot j}$, translates to $\mathbb{E}_j^{-\Omega^{-1/2}}$ which can be applied to functions of ξ. Now we can write the CME (2.39) so that the right side is a function of ξ only:

$$\frac{\partial}{\partial t}P^c(n,t) = \Omega\sum_{j=1}^{r}\left(\mathbb{E}_j^{-\Omega^{-1/2}} - 1\right)\Big[v_j\left(\phi + \Omega^{-1/2}\xi\right) + \mathcal{O}\left(\Omega^{-1}\right)\Big]\Pi(\xi,t), \tag{2.45}$$

where the replacement of $P^c(n,t)$ with $\Pi(\xi,t)$ on the right follows from (2.43). The next step is the Taylor expansion, around ϕ, of $v_j(x)$ and the operator $\mathbb{E}_j^{-\Omega^{-1/2}}$ in several dimensions:

$$v_j\left(\phi + \Omega^{-1/2}\xi\right) = v_j(\phi) + \Omega^{-1/2}\sum_i \frac{\partial v_j}{\partial\phi_i}\xi_i + \mathcal{O}\left(\Omega^{-1}\right),$$

$$\mathbb{E}_j^{-\Omega^{-1/2}} = 1 - \Omega^{-1/2}\sum_i S_{ij}\frac{\partial}{\partial\xi_i} + \frac{1}{2}\Omega^{-1}\sum_{i,k} S_{ij}S_{kj}\frac{\partial^2}{\partial\xi_i\partial\xi_k} + \mathcal{O}\left(\Omega^{-3/2}\right),$$

where the later follows from (2.40) by replacing n with ξ and S with $\Omega^{-1/2}S$. Inserting the above two expansions in (2.45) and then comparing the result with (2.44) leads to

$$\frac{\partial}{\partial t}\Pi(\xi,t) - \Omega^{1/2}\sum_{i=1}^{s}\frac{\mathrm{d}\phi_i}{\mathrm{d}t}\frac{\partial}{\partial\xi_i}\Pi(\xi,t) =$$
$$\Omega\sum_{j=1}^{r}\left(-\Omega^{-1/2}\sum_i S_{ij}\frac{\partial}{\partial\xi_i} + \frac{1}{2}\Omega^{-1}\sum_{i,k} S_{ij}S_{kj}\frac{\partial^2}{\partial\xi_i\partial\xi_k} + \mathcal{O}\left(\Omega^{-3/2}\right)\right)$$
$$\times\left[v_j(\phi) + \Omega^{-1/2}\sum_i\frac{\partial v_j}{\partial\phi_i}\xi_i + \mathcal{O}\left(\Omega^{-1}\right)\right]\Pi(\xi,t).$$

The terms of the order $\Omega^{1/2}$ are proportional to the factors $\partial\Pi/\partial\xi_i$. It is possible to make the terms of each type cancel each other by choosing ϕ such that

$$\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = \sum_{j=1}^{r} S_{ij}v_j(\phi), \tag{2.46}$$

and we see that the macroscopic law emerges as the lowest approximation in the Ω expansion. Comparing terms of the order Ω^0 :

$$\frac{\partial\Pi}{\partial t} = \sum_{j=1}^{r}\left(-\sum_{i,k} S_{ij}\frac{\partial v_j}{\partial\phi_k}\frac{\partial\left(\xi_k\Pi\right)}{\partial\xi_i} + \frac{1}{2}v_j(\phi)\sum_{i,k} S_{ij}S_{kj}\frac{\partial^2\Pi}{\partial\xi_i\partial\xi_k}\right).$$

Introducing matrices A and B with elements

$$A_{ik} = \sum_{j=1}^{r} S_{ij}\frac{\partial v_j}{\partial\phi_k}, \quad \text{and} \quad B_{ik} = \sum_{j=1}^{r} S_{ij}S_{kj}v_j(\phi), \tag{2.47}$$

the above differential equation can be written as

$$\frac{\partial \Pi}{\partial t} = -\sum_{i,k} A_{ik} \frac{\partial \left(\xi_k \Pi\right)}{\partial \xi_i} + \frac{1}{2} \sum_{i,k} B_{ik} \frac{\partial^2 \Pi}{\partial \xi_i \partial \xi_k}, \tag{2.48}$$

which is a linear Fokker-Planck equation (FPE) with coefficient matrices A and B that depend on time through the deterministic reaction rate $v_j\left(\phi(t)\right)$. The solution of the linear FPE (2.48) can be shown to be a multivariate Gaussian

$$\Pi(\xi, t) = (2\pi)^{-s/2} \left|Q(t)\right|^{-1/2} \exp\left(-\frac{1}{2} \xi^T \left[Q(t)\right]^{-1} \xi\right), \tag{2.49}$$

with zero mean and covariance matrix $Q(t)$ that satisfies the matrix *Riccati equation*:

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = AQ + QA^T + B\,. \tag{2.50}$$

The proposed transformation (2.42) together with (2.46) and (2.48) forms the so-called *linear noise approximation* (LNA) [Kam07a].

Since the LNA does not include terms of order higher than Ω^0, the same could have been obtained by applying the method of Ω-expansion to the nonlinear FPE (2.41).

2.6.3 Chemical Langevin equation

This section is based on Gillespie's method [Gil00]. Suppose it is known that $N(t) = n$. During the next short time interval $[t, t+\tau]$, the component-wise copy number will change, according to (2.2), from n_i to

$$N_i(t+\tau) = n_i + \sum_{j=1}^{r} \Big[Z_j(t+\tau) - Z_j(t)\Big] S_{ij}\,. \tag{2.51}$$

Assume the interval length τ is: (i) small enough that each reaction propensity $a_j(n)$ does not change "appreciably" during the interval, and (ii) large enough that the expected number of occurrences $\langle Z_j(t+\tau) - Z_j(t)\rangle$ of each reaction channel R_j during the interval is very large. Assumption (i) allows the Poissonian approximation

$$Z_j(t+\tau) - Z_j(t) \sim \mathcal{P}_j\Big(a_j(n)\,\tau\Big) \tag{2.52}$$

with mean and variance $a_j(n)\,\tau$ for each reaction channel, where all the Poisson processes $\mathcal{P}_j$ are statistically independent. Assumption (ii) implies

$$\langle Z_j(t+\tau) - Z_j(t)\rangle = a_j(n)\,\tau \gg 1,$$

and thus allows the normal approximation

$$\mathcal{P}_j\Big(a_j(n)\tau\Big) \sim \mathcal{N}_j\Big(a_j(n)\,\tau, a_j(n)\,\tau\Big) \tag{2.53}$$

with the same mean and variance $a_j(n)\tau$ for each reaction channel, where all the normal processes $\mathcal{N}_j$ are statistically independent. Since any normal variable can be expressed as

a sum of its mean and the standard normal variable scaled by its standard deviation, we have

$$\mathcal{N}_j\Big(a_j(n)\tau, a_j(n)\tau\Big) = a_j(n)\tau + (a_j(n)\tau)^{1/2}\,\mathcal{N}_j(0,1)\,. \tag{2.54}$$

Backward substitution through (2.51)-(2.54) leads to continuous approximation

$$N_i^c(t+\tau) = n + \sum_{j=1}^{r} S_{ij} a_j(n)\tau + \sum_{j=1}^{r} S_{ij}\,(a_j(n)\tau)^{1/2}\,\mathcal{N}_j(0,1),$$

where $N^c(t)$ denotes the continuous Markov process approximating the jump process $N(t)$. The factor $\sqrt{\tau}\mathcal{N}_j(0,1)$ in the 2nd summation on the right suggests a set $\{\mathcal{W}_j(t)\}$ of independent standard *Brownian motions*, or standard *Wiener processes*, with Wiener increments,

$$\mathcal{W}_j(t+\tau) - \mathcal{W}_j(t) = \sqrt{\tau}\mathcal{N}_j(0,1)\,.$$

Combining this with $N(t) = n$, the previous equation can be written as

$$N_i^c(t+\tau) - N_i^c(t) = \sum_{j=1}^{r} S_{ij} a_j\left(N^c(t)\right)\tau + \sum_{j=1}^{r} S_{ij}\Big(a_j\left(N^c(t)\right)\Big)^{1/2}\left[\mathcal{W}_j(t+\tau) - \mathcal{W}_j(t)\right]\,.$$

Making the replacement $\tau = dt$ gives the stochastic differential equation

$$dN_i^c(t) = \sum_{j=1}^{r} S_{ij} a_j\left(N^c(t)\right) dt + \sum_{j=1}^{r} S_{ij}\Big(a_j\left(N^c(t)\right)\Big)^{1/2} d\mathcal{W}_j(t), \tag{2.55}$$

known as the "standard-form" *chemical Langevin equation* (CLE). An equivalent "white-noise form" of CLE can be written as

$$\frac{\mathrm{d}N_i^c(t)}{\mathrm{d}t} = \sum_{j=1}^{r} S_{ij} a_j\left(N^c(t)\right) + \sum_{j=1}^{r} S_{ij}\Big(a_j\left(N^c(t)\right)\Big)^{1/2}\Gamma_j(t), \tag{2.56}$$

where $\Gamma_j(t) = {}^{\mathrm{d}W_j}/_{\mathrm{d}t}$ are statistically independent Gaussian white-noise processes.

It was shown in [Gil96] that the probability density function $P^c(n,t)$ of the continuous Markov process $N^c(t)$ obeys the FPE (2.41). Thus the CLE (2.55) and (2.41) are equivalent.

The two conditions (i) and (ii) seem conflicting and require the existence of a domain of macroscopically infinitesimal time intervals. Although the existence of a such a domain cannot be guaranteed, Gillespie argues that this can be found for most practical cases. Admitting that, "it may not be easy to continually monitor the system to ensure that conditions (i) and (ii) [..] are satisfied." He justifies his argument by saying that this "will not be the first time that Nature has proved to be unaccommodating to our purposes." [Gil00].

Generating sample paths of (2.56) is orders of magnitude faster than doing the same for the CME because it essentially needs generation of normal random numbers. See [Hig01] for numerical simulation methods of stochastic differential equations such as (2.56). However, solving the nonlinear FPE (2.41) for the probability density is as difficult as the CME. Therefore, from an analytical point of view, the CLE and the associated nonlinear

FPE do not provide any significant advantage. However, linearising the propensity function around the mean [Gou06], or using the Ω-expansion [Kam07a], the nonlinear FPE (2.41) can be reduced to the LNA whose solution is a Gaussian distribution with a mean that is equal to the solution of the deterministic ODE model and a covariance matrix that obeys a linear ODE. This is the main drawback of LNA because, for system containing at least one biomolecular reactions, the mean of a stochastic model is not equal to the solution of deterministic ODEs, as shown next.

2.7 Markov processes with both continuous and jump character

The CME describes a jump process while the FPE describes a diffusion process. A general Markov process in continuous time can be made up of both continuous and jump parts. Consider a Markov process $Y(t)$ taking real values y from a state space $\mathcal{Y}$ with probability densities $p(y,t)$. If the process has only continuous sample paths, changes in $p(y,t)$ are described by the FPE. If the process has only discontinuous sample paths (with jumps), changes in the density will be governed by a so called master equation:

$$\frac{\partial}{\partial t}p(y,t) = \int_{\mathcal{Y}} \Big[W(y \,|\, y')p(y',t) - W(y' \,|\, y)p(y,t)\Big]dy',$$

a generalisation of the CME where the summation has been replaced by integration (because of the real state space) and the propensity by the general transition rate $W(y \,|\, q)$ defined for the transitions $q \to y$. If the process has sample paths made up of both the continuous and jump parts, changes in the density $p(y,t)$ will be governed by a differential equation obtained by combining the above master equation and the FPE to get what is known as the *differential Chapman Kolmogorov equation* (dCKE):

$$\frac{\partial}{\partial t}p(y,t) = -\sum_i \frac{\partial}{\partial y_i}\Big[\alpha_i(y)p(y,t)\Big] + \frac{1}{2}\sum_{i,k}\frac{\partial^2}{\partial y_i \partial y_k}\Big[\beta_{ik}(y)p(y,t)\Big] + \int_{\mathcal{Y}} \Big[W(y \,|\, y')p(y',t) - W(y' \,|\, y)p(y,t)\Big]dy' .$$

For detailed derivation, see [UW07, Gar04]. The dCKE describes a variety of Markov processes: It becomes an FPE for $W = 0$, a master equation for $\alpha = 0$, $\beta = 0$, a Liouville equation for $\beta = 0$, $W = 0$ and a Liouville master equation for $\beta = 0$. The way in which different equations and the corresponding models stem from the dCKE is sketched in Figure 2.2 which also shows how simulation strategies connect to their modelling counterparts. For a detailed discussion, see [Gar04] and our work in [UW07].

The analytical approximations discussed in this chapter do not allow direct tracking of the mean and (co)variance which, in general, are coupled. This coupling is not obvious in CLE and FPE, and ignored by LNA and conventional ODE models. The next chapter presents the 2MA approach, which has a direct representation of the first two moments and the coupling between them.

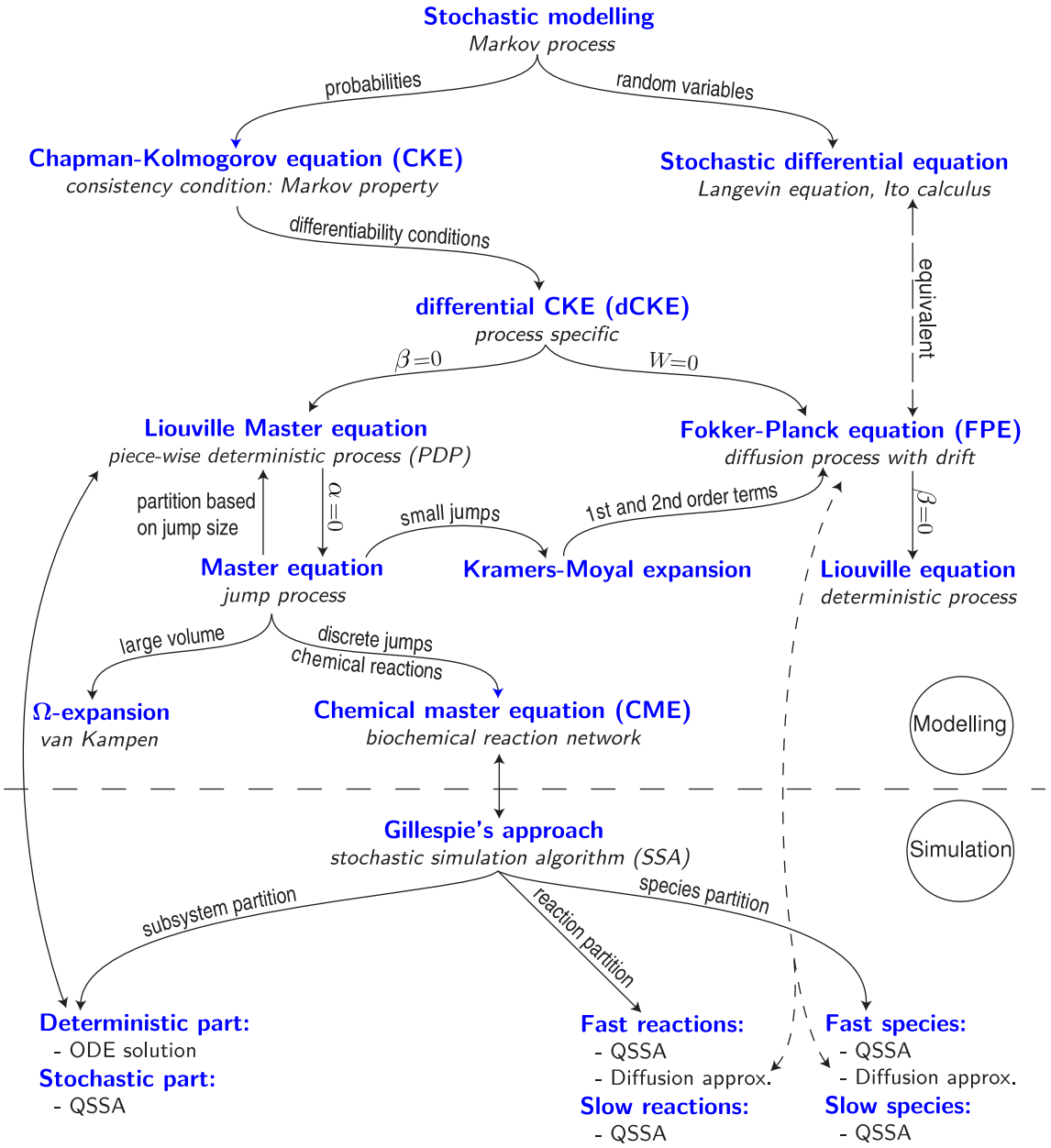

Figure 2.2 Interrelationships for various stochastic approaches. The coefficients α, β, W respectively refer to the drift-vector, diffusion-matrix and the transition-rate in the dCKE. QSSA stands for Quasi-Steady-State Assumption. For details, see [UW07].

3

The 2MA approach

This chapter develops a compact form of the 2MA equations - a system of ODEs for the dynamics of the mean and (co)variance of the continuous-time discrete-state Markov process that models a biochemical reaction system by the CME. This is an extension of previous derivations, taking into account relative concentrations and non-elementary reactions. The compact form, obtained by careful selection of notation, allows for an easy interpretation.

3.1 Relative concentrations

The concentration $X_i(t)$ is, in general, the copy number $N_i(t)$ divided by some fixed scaling parameter Ω_i specific to that component. In other words

$$N_i(t) = \Omega_i X_i(t)\,.$$

The copy number and concentration (vectors) are related by

$$N(t) = \Omega \odot X(t),$$

where Ω is the s-vector of elements Ω_i and $\odot$ denotes the element-wise product of two vectors (or matrices). Note that our approach is developed for the general case which allows for relative concentrations instead of assuming one global system-size Ω.

Often we are interested in the first two moments of the probability distribution. The first moment is the mean vector $\langle N(t)\rangle$ of copy numbers, defined element-wise by

$$\langle N_i(t)\rangle = \sum_n n_i P(n,t),$$

the ith mean copy number. The 2nd central moment is the covariance matrix $\left\langle N, N^T\right\rangle$ defined element-wise by

$$\langle N_i, N_k\rangle = \Big\langle \Big(N_i - \langle N_i\rangle\Big)\Big(N_k - \langle N_k\rangle\Big)\Big\rangle,$$

the covariance between N_i and N_k. Here the superscript T denotes transpose of a matrix. When obvious from the context, we will leave out dependence on time, as in the above.

We are also interested in the mean concentration vector $\langle X \rangle$ with elements

$$\langle X_i \rangle = \frac{\langle N_i \rangle}{\Omega_i}$$

and the concentration covariance matrix $\left\langle X, X^T \right\rangle$ with elements

$$\langle X_i, X_k \rangle = \frac{\langle N_i, N_k \rangle}{\Omega_i \Omega_k} .$$

The diagonal elements of the covariance matrix are the variances:

$$\mathrm{Var}\Big[N_i\Big] = \langle N_i, N_i \rangle , \quad \mathrm{Var}\Big[X_i\Big] = \langle X_i, X_i \rangle .$$

3.2 Dynamics of the mean

Taking expectation on both sides of (2.2) gives the mean copy number,

$$\langle N_i(t) \rangle = N_i(0) + \sum_{j=1}^{r} S_{ij} \langle Z_j(t) \rangle .$$

Taking the time-derivative and employing (2.25) yields

$$\frac{\mathrm{d} \langle N_i \rangle}{\mathrm{d}t} = \sum_{j=1}^{r} S_{ij} \langle a_j(N) \rangle = \langle f_i(N) \rangle , \tag{3.1}$$

where

$$f_i(n) \stackrel{\text{def}}{=} \sum_{j=1}^{r} S_{ij} a_j(n) \tag{3.2}$$

referred to here as the *copy-number flux*. Dividing by Ω_i gives the system of ODEs for the mean concentration:

$$\frac{\mathrm{d} \langle X_i \rangle}{\mathrm{d}t} = \left\langle \tilde{f}_i(X) \right\rangle \tag{3.3}$$

where

$$\tilde{f}_i(x) \stackrel{\text{def}}{=} \frac{f_i(\Omega \odot x)}{\Omega_i} = \frac{1}{\Omega_i} \sum_{j=1}^{r} S_{ij} a_j(\Omega \odot x) \tag{3.4}$$

referred to here as the *concentration flux*. When Ω is a scalar (system size), then (2.28) allows a simpler expression for concentration flux

$$\tilde{f}(x) = \frac{f(\Omega x)}{\Omega} = S\, \tilde{a}(x) . \tag{3.5}$$

It is interesting to note that (3.1) is a direct consequence of mass conservation (2.2) and definition of propensity because we have not referred to the CME (which is the usual procedure) during our derivation.

Isomerisation revisited: Following from the state transitions (2.31) for the isomerisation reaction scheme (2.5), the reaction propensities are linear

$$a_1 = k_1 n, \quad a_2 = \Big(n^{\text{tot}} - n\Big)\, k_2,$$

giving the linear copy-number flux

$$f(n) = -a_1 + a_2 = k_2 n^{\text{tot}} - (k_1 + k_2)\, n\,.$$

The mean copy number thus obeys

$$\frac{\mathrm{d}\,\langle N\rangle}{\mathrm{d}t} = \big\langle f(N)\big\rangle = k_2 n^{\text{tot}} - (k_1 + k_2)\,\langle N\rangle\,,$$

which is the same as the deterministic ODE (2.13). In general, the mean of a system composed solely of reactions of zero and/or first-order reactions is the same as the solution of the corresponding deterministic ODE because of linear propensities.

Lotka-Volterra model revisited: Following the state transitions (2.32) for the predator-prey reaction scheme (2.6), the reaction propensities are given by

$$a_1 = k_1 n_{\text{A}} n_1, \quad a_2 = k_2 n_1 n_2, \quad a_3 = k_3 n_2,$$

giving the copy number flux vector

$$f(n) = S\,a(n) = \begin{bmatrix} a_1 - a_2 \\ \\ a_2 - a_3 \end{bmatrix} = \begin{bmatrix} k_1 n_{\text{A}} n_1 - k_2 n_1 n_2 \\ \\ k_2 n_1 n_2 - k_3 n_2 \end{bmatrix}\,.$$

The mean copy number changes according to

$$\begin{aligned} \frac{\mathrm{d}\,\langle N_1\rangle}{\mathrm{d}t} &= \big\langle f_1(N)\big\rangle = k_1 n_{\text{A}}\,\langle N_1\rangle - k_2\,\langle N_1 N_2\rangle\,, \\ \frac{\mathrm{d}\,\langle N_2\rangle}{\mathrm{d}t} &= \big\langle f_2(N)\big\rangle = k_2\,\langle N_1 N_2\rangle - k_3\,\langle N_2\rangle\ . \end{aligned}$$

Since $\langle N_1 N_2\rangle \neq \langle N_1\rangle\,\langle N_2\rangle$, the mean $\langle N\rangle$ is not the same as the solution of the deterministic ODE (2.14). In general, the mean of a system containing 2nd and/or higher-order reactions is not the same as the solution of the corresponding deterministic ODE because of nonlinear propensities. Hence the deterministic model cannot be used, in general, to describe the correct mean.

In general, the mean flux $\big\langle f(N)\big\rangle$ involves the unknown probability distribution $P(n,t)$. In other words, the mean copy number depends not just on the mean itself, but also involves higher-order moments, and therefore (3.3) is, in general, not closed in $\langle N\rangle$. Suppose the propensity $a_j(n)$ is a smooth function and that central moments $\langle (N - \langle N\rangle)^m\rangle$ of order higher than $m = 2$ can be ignored. In that case the Taylor series expansion of copy-number flux $f_i(n)$ around the mean $\langle N\rangle$ is

$$f_i(n) = f_i\Big(\langle N\rangle\Big) + \frac{\partial f_i}{\partial n^T}\Big(n - \langle N\rangle\Big) + \frac{1}{2}\Big(n - \langle N\rangle\Big)^T \frac{\partial^2 f_i}{\partial n \partial n^T}\Big(n - \langle N\rangle\Big) + \cdots\,.$$

All the partial derivatives with respect to the state n are evaluated at $n = \langle N \rangle$. The first-order partial derivative here is the ith row of the Jacobian $\frac{\partial f}{\partial n^T}$. The second-order partial derivative is the Hessian of f_i. Expectation of the 2nd term on the right is zero. Ignoring terms (moments) of order higher than two, the ODE (3.1) can be approximated by:

$$\frac{\mathrm{d}\langle N_i \rangle}{\mathrm{d}t} = f_i\Big(\langle N \rangle\Big) + \frac{1}{2}\frac{\partial^2 f_i}{\partial n \partial n^T} : \Big\langle N, N^T \Big\rangle \,. \tag{3.6}$$

Here $A : B$ denotes the Frobenius inner product, the sum of products of the corresponding elements, between the two matrices A and B. The last term on the right, referred to here as the *stochastic copy-number flux*, arises from the discrete and random nature of chemical reactions. Note that this term has been derived from the CME instead of being assumed like external noise. This shows that knowledge of fluctuations (even if small) is important for a correct description of the mean. This also indicates an advantage of the stochastic framework over its deterministic counterpart: starting from the same assumptions and approximations, the stochastic framework allows us to see the influence of fluctuation on the mean. Note that the above equation is exact for systems where no reaction has an order higher than two because then 3rd and higher derivatives of propensity are zero.

Repeating the above procedure of Taylor series expansion for the concentration flux $\tilde{f}_i(x)$ around the mean $\langle X \rangle$ will lead to the analogous system of ODEs

$$\frac{\mathrm{d}\langle X_i \rangle}{\mathrm{d}t} = \tilde{f}\Big(\langle X \rangle\Big) + \frac{1}{2}\frac{\partial^2 \tilde{f}}{\partial x \partial x^T} : \Big\langle X, X^T \Big\rangle \,, \tag{3.7}$$

where all the partial derivatives with respect to the state x are evaluated at $x = \langle X \rangle$.

3.3 Dynamics of the (co)variance

Before we can see how the (co)variances $\langle N_i N_k \rangle$ evolve in time, let us multiply the CME with $n_i n_k$ and sum over all n,

$$\begin{aligned}
\sum_n n_i n_k \frac{\mathrm{d}P(n,t)}{\mathrm{d}t} &= \sum_n n_i n_k \sum_{j=1}^{r} \Big[a_j(n - S_{\bullet j}) P(n - S_{\bullet j}, t) - a_j(n) P(n,t) \Big] \\
&= \sum_n \sum_{j=1}^{r} \Big[(n_i + S_{ij})\,(n_k + S_{kj})\, a_j(n) P(n,t) - n_i n_k a_j(n) P(n,t) \Big] \\
&= \sum_n \sum_{j=1}^{r} \Big(n_k S_{ij} + n_i S_{kj} + S_{ij} S_{kj} \Big) a_j(n) P(n,t)
\end{aligned}$$

where dependence on time is implicit for all variables except n and s. Recognising sums of probabilities as expectations,

$$\frac{\mathrm{d}\langle N_i N_k \rangle}{\mathrm{d}t} = \langle N_k f_i(N) \rangle + \langle N_i f_k(N) \rangle + \langle B_{ik}(N) \rangle$$

where

$$B_{ik}(n) \stackrel{\text{def}}{=} \sum_{j=1}^{r} S_{ij} S_{kj} a_j(n) \tag{3.8}$$

forms the (i,k)th element of the $s \times s$ *diffusion matrix* $B(n)$. The relation

$$\left\langle N, N^T \right\rangle = \left\langle N N^T \right\rangle - \langle N \rangle \langle N \rangle^T$$

can be utilised to yield

$$\frac{\mathrm{d}\langle N_i, N_k \rangle}{\mathrm{d}t} = \langle (N_k - \langle N_k \rangle) f_i(N) \rangle + \langle (N_i - \langle N_i \rangle) f_k(N) \rangle + \langle B_{ik}(N) \rangle \tag{3.9}$$

for the copy-number covariance. Multiplying by the inverse of $\Omega_i \Omega_k$ gives the analogous system of ODEs for concentration covariance:

$$\frac{\mathrm{d}\langle X_i, X_k \rangle}{\mathrm{d}t} = \left\langle (X_k - \langle X_k \rangle) \tilde{f}_i(X) \right\rangle + \left\langle (X_i - \langle X_i \rangle) \tilde{f}_k(X) \right\rangle + \frac{\left\langle \tilde{B}_{ik}(X) \right\rangle}{\sqrt{\Omega_i \Omega_k}} \tag{3.10}$$

where the diffusion matrix $\tilde{B}$ has (i,k)th element

$$\tilde{B}_{ik}(x) \stackrel{\text{def}}{=} \frac{B_{ik}(\Omega \odot x)}{\sqrt{\Omega_i \Omega_k}} = \frac{1}{\sqrt{\Omega_i \Omega_k}} \sum_{j=1}^{r} S_{ij} S_{kj} a_j(\Omega \odot x)\,. \tag{3.11}$$

When Ω is a scalar (system size), (2.28) allows a simpler expression for the above

$$\tilde{B}_{ik}(x) = \frac{B_{ik}(\Omega x)}{\Omega} = \sum_{j=1}^{r} S_{ij} S_{kj} \tilde{a}_j(x)\,. \tag{3.12}$$

Let us start with the component form of (3.9). The argument of the first expectation on the right has Taylor expansion

$$f_i(n)\Big(n_k - \langle N_k \rangle\Big) = f_i\Big(\langle N \rangle\Big)\Big(n_k - \langle N_k \rangle\Big) + \frac{\partial f_i}{\partial n^T}(n - \langle N \rangle)\Big(n_k - \langle N_k \rangle\Big) + \cdots\,.$$

Expectation of the first term on the right is zero. Ignoring moments of order higher than two, the first expectation in (3.9) is then

$$\left\langle (N_k - \langle N_k \rangle) f_i(N) \right\rangle = \frac{\partial f_i}{\partial n^T} \langle N, N_k \rangle\ .$$

By a similar procedure, the second expectation in (3.9) is

$$\left\langle \Big(N_i - \mu_{N,i}\Big) f_k(N) \right\rangle = \left\langle N_i, N^T \right\rangle \frac{\partial f_k}{\partial n},$$

correct to 2nd-order moments. The element $B_{ik}(n)$ of the diffusion matrix has Taylor expansion

$$B_{ik}(n) = B_{ik}\Big(\langle N \rangle\Big) + \frac{\partial B_{ik}}{\partial n^T}\Big(n - \langle N \rangle\Big) + \frac{1}{2}\Big(n - \langle N \rangle\Big)^T \frac{\partial^2 B_{ik}}{\partial n \partial n^T}\Big(n - \langle N \rangle\Big) + \cdots\,.$$

Taking term-wise expectation, and ignoring 3rd and higher-order moments,

$$\langle B_{ik}(N) \rangle = B_{ik}\Big(\langle N \rangle\Big) + \frac{1}{2} \frac{\partial^2 B_{ik}}{\partial n \partial n^T} : \left\langle N, N^T \right\rangle\ .$$

Summing up the three expectations above gives the ODE

$$\frac{\mathrm{d}\,\langle N_i, N_k\rangle}{\mathrm{d}t} = \frac{\partial f_i}{\partial n^T}\,\langle N, N_k\rangle + \Big\langle N_i, N^T\Big\rangle\,\frac{\partial f_k}{\partial n} + B_{ik}\Big(\langle N\rangle\Big) + \frac{1}{2}\frac{\partial^2 B_{ik}}{\partial n \partial n^T} : \Big\langle N, N^T\Big\rangle \quad (3.13)$$

for the component-wise covariances. The drift matrix $^{\partial f}/_{\partial n^T}$ reflects the dynamics for relaxation (dissipation) to the steady state and the the (Taylor approximation to the 2nd-order of) diffusion matrix B the randomness (fluctuation) of the individual events [PE06]. These terms are borrowed from the *fluctuation-dissipation theorem* (FDT) [Kei87, Lax60] which has the same form as (4.2). Remember that (4.2) is exact for systems that contain only zero- and first-order reactions because in that case the propensity is already linear.

Repeating the above procedures of Taylor series expansions around the mean $\langle X\rangle$ for the analogous terms to the right of (3.10) will lead to the analogous system of ODEs

$$\frac{\mathrm{d}\,\langle X_i, X_k\rangle}{\mathrm{d}t} = \frac{\partial \tilde{f}_i}{\partial x^T}\,\langle X, X_k\rangle + \Big\langle X_i, X^T\Big\rangle\,\frac{\partial \tilde{f}_k}{\partial x} + \frac{1}{\sqrt{\Omega_i \Omega_k}}\left[\tilde{B}_{ik}\Big(\langle X\rangle\Big) + \frac{1}{2}\frac{\partial^2 \tilde{B}_{ik}}{\partial x \partial x^T} : \Big\langle X, X^T\Big\rangle\right], \quad (3.14)$$

for the pair-wise concentration covariance.

3.4 Outline of the 2MA and examples

We can summarise the 2MA method, in terms of copy numbers, by the following steps:

1. Assign propensity $a_j(n)$ to each reaction channel.
2. Construct the flux $f_i(n)$ according to(3.2) and the partial derivatives $\frac{\partial f_i}{\partial n^T}$ and $\frac{\partial^2 f_i}{\partial n \partial n^T}$, both evaluated at $n = \langle N\rangle$, for each species.
3. Construct the diffusion coefficient $B_{ik}(n)$ according to (3.8) and the partial derivative $\frac{\partial^2 B_{ik}}{\partial n \partial n^T}$, evaluated at $n = \langle N\rangle$, for each pair of species.
4. Construct the element-wise product $\frac{\partial^2 f_i}{\partial n \partial n^T} : \langle N, N^T\rangle$ between the Hessian of f_i and covariance matrix, for each species.
5. Construct the element-wise product $\frac{\partial^2 B_{ik}}{\partial n \partial n^T} : \langle N, N^T\rangle$ between the Hessian of B_{ik} and covariance matrix for each pair of species.
6. Insert the above expressions in (3.6) and (3.13) to obtain the 2MA equations.

The procedure for the 2MA equations in terms of concentrations is the same except for appropriate replacements such as N with X and so on.

Isomerisation revisited: Following from the state transitions (2.31) for the isomerisation reaction scheme (2.5), the reaction propensities are

$$a_1 = k_1 n, \quad a_2 = \left(n^{\text{tot}} - n\right) k_2 \,.$$

The copy-number flux is given by

$$f(n) = S\, a(n) = -k_1 n + \left(n^{\text{tot}} - n\right) k_2,$$

and the corresponding diffusion coefficient by

$$B(n) = k_1 n + \left(n^{\text{tot}} - n\right) k_2 \,.$$

The drift matrix is given by

$$\frac{\partial f}{\partial n} = -\left(k_1 + k_2\right) .$$

The Hessians of f and B are both zero:

$$\frac{\partial^2 f}{\partial n^2} = 0, \quad \frac{\partial^2 B}{\partial n^2} = 0 \,.$$

Finally we obtain the 2MA equations:

$$\begin{aligned} \frac{\mathrm{d}\,\langle N\rangle}{\mathrm{d}t} &= -\left(k_1 + k_2\right)\langle N\rangle + k_2 n^{\text{tot}}, \\ \frac{\mathrm{d}\,\mathrm{Var}\,[N]}{\mathrm{d}t} &= -2\left(k_1 + k_2\right)\mathrm{Var}\,[N] + \left(k_1 - k_2\right)\langle N\rangle + k_2 n^{\text{tot}} \,. \end{aligned} \tag{3.15}$$

We see that the growth of variance is influenced by the mean through the rate term. With a rise in the mean, the growth of variance speeds up if $k_1 > k_2$, slows down if $k_1 < k_2$ and is not influenced if $k_1 = k_2$. This is illustrated in Figure 3.1 which plots the standard deviation (std) and the coefficient of variation (CV) for three pairs of parameter values with the same sum $k_1 + k_2 = 4$. It is interesting to note that the transient overshoot of the std is not shared by the CV. To get a qualitative idea about possible stochastic realisations, the mean and the band of one std around it are plotted in Figure 3.2 for the same three pairs of parameter values. In the non-dimensional time, $\tau = (k_1 + k_2)t$, the above pair of ODEs takes the form

$$\begin{aligned} \frac{\mathrm{d}\,\langle N\rangle}{\mathrm{d}\tau} &= -\langle N\rangle + \left(\frac{k_2}{k_1 + k_2}\right) n^{\text{tot}}, \\ \frac{\mathrm{d}\,\mathrm{Var}\,[N]}{\mathrm{d}\tau} &= -2\,\mathrm{Var}\,[N] + \left(\frac{k_1 - k_2}{k_1 + k_2}\right)\langle N\rangle + \left(\frac{k_2}{k_1 + k_2}\right) n^{\text{tot}} \,. \end{aligned}$$

Now we can see that experimental data on both the mean and variance is needed for identifiability of both the parameters. The mean alone gives information about the fraction $k_2/(k_1+k_2)$ only.

Here we have used the std and CV as measures of noise. In addition to these two, other alternative measures of noise, including the Fano factor $F = \mathrm{Var}[N]/\langle N\rangle$ and the noise-to-signal ratio $\zeta = \mathrm{Var}[N]/\langle N\rangle^2$, are discussed in [Pau05].

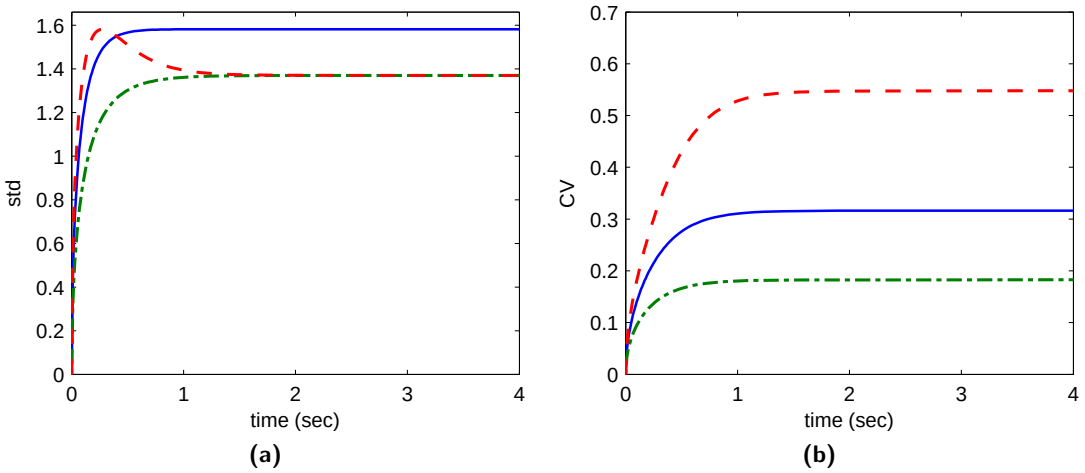

Figure 3.1 The time courses, for the isomerisation reaction, of : (a) standard deviation (std), (b) coefficient of variation (CV), for the parameter pairs $k_1 = k_2 = 2\,\mathrm{sec}^{-1}$ as solid lines, $k_1 = 1\,\mathrm{sec}^{-1}$, $k_2 = 3\,\mathrm{sec}^{-1}$ in dash-dotted lines, and $k_1 = 3\,\mathrm{sec}^{-1}$, $k_2 = 1\,\mathrm{sec}^{-1}$ in dashed lines. All the parameter values satisfy $k_1 + k_2 = 4$.

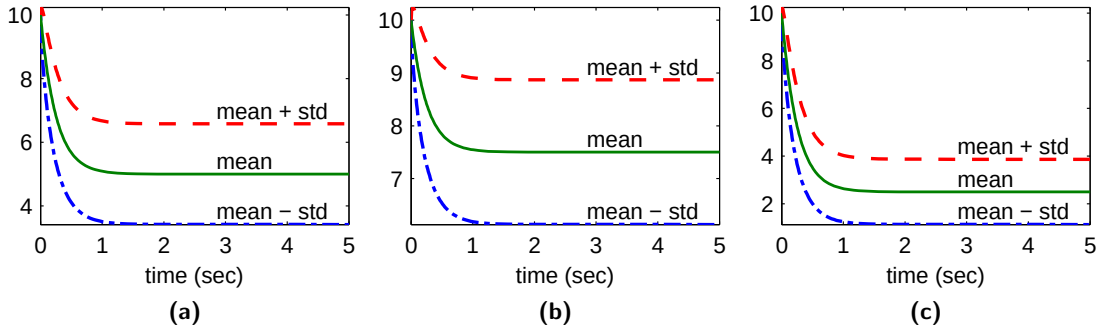

Figure 3.2 The time courses of mean and mean±std for the isomerisation reaction when the parameters are: (a) $k_1 = k_2 = 2\,\mathrm{sec}^{-1}$, (b) $k_1 = 1\,\mathrm{sec}^{-1}$, $k_2 = 3\,\mathrm{sec}^{-1}$, and (c) $k_1 = 3\,\mathrm{sec}^{-1}$, $k_2 = 1\,\mathrm{sec}^{-1}$. All the parameter values satisfy $k_1 + k_2 = 4$.

Lotka-Volterra model revisited: Following the state transitions (2.32) for the predator-prey reaction scheme (2.6), the reaction propensities are given by

$$a_1 = k_1 n_{\mathrm{A}} n_1, \quad a_2 = k_2 n_1 n_2, \quad a_3 = k_3 n_2 \,.$$

The copy number flux is given by

$$f(n) = S\, a(n) = \begin{bmatrix} a_1 - a_2 \\ \\ a_2 - a_3 \end{bmatrix} = \begin{bmatrix} k_1 n_{\mathrm{A}} n_1 - k_2 n_1 n_2 \\ \\ k_2 n_1 n_2 - k_3 n_2 \end{bmatrix},$$

and the corresponding diffusion matrix by

$$B = \begin{bmatrix} a_1 - a_2 & -a_2 \\ \\ -a_2 & a_2 - a_3 \end{bmatrix} = \begin{bmatrix} k_1 n_{\mathrm{A}} n_1 - k_2 n_1 n_2 & -k_2 n_1 n_2 \\ \\ -k_2 n_1 n_2 & k_2 n_1 n_2 - k_3 n_2 \end{bmatrix}.$$

The drift matrix can be worked out to be

$$\frac{\partial f}{\partial n^T} = \begin{bmatrix} k_1 n_{\mathrm{A}} - k_2 \langle N_2 \rangle & -k_2 \langle N_1 \rangle \\ \\ k_2 \langle N_2 \rangle & k_2 \langle N_1 \rangle - k_3 \end{bmatrix}.$$

The Hessian matrices of the elements of f can be worked out to be

$$\frac{\partial^2 f_1}{\partial n \partial n^T} = \begin{bmatrix} 0 & -k_2 \\ \\ -k_2 & 0 \end{bmatrix}, \quad \frac{\partial^2 f_2}{\partial n \partial n^T} = \begin{bmatrix} 0 & k_2 \\ \\ k_2 & 0 \end{bmatrix}.$$

The Hessian matrices of the elements of B can be worked out to be

$$\frac{\partial^2 B_{11}}{\partial n \partial n^T} = \frac{\partial^2 B_{12}}{\partial n \partial n^T} = \frac{\partial^2 B_{21}}{\partial n \partial n^T} = \begin{bmatrix} 0 & -k_2 \\ \\ -k_2 & 0 \end{bmatrix}, \quad \frac{\partial^2 B_{22}}{\partial n \partial n^T} = \begin{bmatrix} 0 & k_2 \\ \\ k_2 & 0 \end{bmatrix}.$$

The above expressions for f and its Hessian can now be inserted in (3.6) to obtain the first set of the 2MA equations: ODEs for the concentration mean. The expressions for B and its Hessian, together with the drift matrix $\partial f / \partial n^T$ can be (3.13) to obtain the 2nd set of 2MA equations: ODEs for the concentration (co)variance matrix.

Enzyme kinetic model revisited: Following the state transitions (2.33) for the enzyme kinetic reaction scheme (2.7), the reaction rates are given by

$$v_1 = \left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}} \right) k_1 x_{\mathrm{S}}, \quad v_2 = k_2 x_{\mathrm{ES}}, \quad v_3 = k_3 x_{\mathrm{ES}} \,.$$

The concentration flux is given by

$$\tilde{f}(x) = S\,\tilde{a}(x) = S\,v(x) = \begin{bmatrix} -\left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} + k_2 x_{\mathrm{ES}} \\ \left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} - k_2 x_{\mathrm{ES}} - k_3 x_{\mathrm{ES}} \end{bmatrix},$$

and the corresponding diffusion matrix by

$$\tilde{B}(x) = \begin{bmatrix} \left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} + k_2 x_{\mathrm{ES}} & -\left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} - k_2 x_{\mathrm{ES}} \\ -\left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} - k_2 x_{\mathrm{ES}} & \left(x_{\mathrm{S}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} + k_2 x_{\mathrm{ES}} + k_3 x_{\mathrm{ES}} \end{bmatrix}.$$

The drift matrix can be worked out to be

$$\frac{\partial \tilde{f}}{\partial x^T} = \begin{bmatrix} -\left(x_{\mathrm{S}}^{\mathrm{tot}} - \left\langle x_{\mathrm{ES}} \right\rangle\right) k_1 & k_1 \left\langle x_{\mathrm{S}} \right\rangle + k_2 \\ \left(x_{\mathrm{S}}^{\mathrm{tot}} - \left\langle x_{\mathrm{ES}} \right\rangle\right) k_1 & -k_1 \left\langle x_{\mathrm{S}} \right\rangle - k_2 - k_3 \end{bmatrix}.$$

The Hessian matrices of the elements of $\tilde{f}$ can be worked out to be

$$\frac{\partial^2 \tilde{f}_1}{\partial x \partial x^T} = \begin{bmatrix} 0 & k_1 \\ k_1 & 0 \end{bmatrix}, \quad \frac{\partial^2 \tilde{f}_2}{\partial x \partial x^T} = \begin{bmatrix} 0 & -k_1 \\ -k_1 & 0 \end{bmatrix}.$$

The Hessian matrices of the elements of $\tilde{B}$ can be worked out to be

$$\frac{\partial^2 \tilde{B}_{11}}{\partial x \partial x^T} = \frac{\partial^2 \tilde{B}_{22}}{\partial x \partial x^T} = \begin{bmatrix} 0 & -k_1 \\ -k_1 & 0 \end{bmatrix}, \quad \frac{\partial^2 \tilde{B}_{12}}{\partial x \partial x^T} = \frac{\partial^2 \tilde{B}_{21}}{\partial x \partial x^T} = \begin{bmatrix} 0 & k_1 \\ k_1 & 0 \end{bmatrix}.$$

The above expressions for $\tilde{f}$ and its Hessian can now be inserted in (3.6) to obtain the first set of the 2MA equations: ODEs for the concentration mean. The expressions for $\tilde{B}$ and its Hessian, together with the drift matrix $\partial \tilde{f} / \partial x^T$ can be (3.13) to obtain the 2nd set of 2MA equations: ODEs for the concentration (co)variance matrix.

Schlögl model revisited: Following the state transitions (2.34) for the Schlögl reaction scheme (2.8), the reaction propensities are given by

$$a_1 = (n-1)\hat{k}_1 n + \hat{k}_3, \quad a_2 = (n-2)(n-1)\hat{k}_2 n + k_4 n,$$

where

$$\hat{k}_1 = \frac{k_1 x_{\mathrm{A}}}{\Omega}, \quad \hat{k}_2 = \frac{k_2}{\Omega^2}, \quad \hat{k}_3 = k_3 x_{\mathrm{B}} \Omega\,.$$

The copy-number flux is given by

$$f = S\,a(n) = -\hat{k}_2 n^3 + \left(\hat{k}_1 + 3\hat{k}_2\right) n^2 - \left(k_4 + 2\hat{k}_2 + \hat{k}_1\right) n + \hat{k}_3,$$

and the corresponding diffusion coefficient by

$$B = \hat{k}_2 n^3 + \left(\hat{k}_1 - 3\hat{k}_2\right) n^2 + \left(k_4 + 2\hat{k}_2 - \hat{k}_1\right) n + \hat{k}_3 \,.$$

The drift coefficient can be worked out to be

$$\frac{\partial f}{\partial n} = -3\hat{k}_2 \langle N \rangle^2 + 2\left(\hat{k}_1 + 3\hat{k}_2\right) \langle N \rangle - \hat{k}_1 - 2\hat{k}_2 - k_4 \,.$$

The Hessian of f can be worked out to be

$$\frac{\partial^2 f}{\partial x^2} = -6\hat{k}_2 \langle N \rangle + 2\hat{k}_1 + 6\hat{k}_2 \,.$$

The Hessian of B can be worked out to be

$$\frac{\partial^2 B}{\partial x^2} = 6\hat{k}_2 \langle N \rangle + 2\hat{k}_1 - 6\hat{k}_2 \,.$$

The above expressions for f and its Hessian can now be inserted in (3.6) to obtain the first set of the 2MA equations: ODEs for the concentration mean. The expressions for B and its Hessian, together with the drift matrix $\partial f/\partial n$ can be (3.13) to obtain the 2nd set of 2MA equations: ODEs for the concentration (co)variance matrix.

The examples used above served as an illustration of the 2MA method. The next chapter investigates the 2MA approach for a practical example of a complex system with non-elementary reactions and relative concentration.

4

The 2MA cell cycle model

This chapter takes the Tyson-Novák model [NPCT01] for the fission yeast cell cycle as a case study. This deterministic model is a practical example using non-elementary reactions and relative concentrations, the two central features of our extended 2MA approach. This will allow us to investigate the price of higher-order truncations by comparing the simulated cycle time statistics with experiments.

4.1 The 2MA equations revisited

In this chapter, we adopt a simplified notation, for the mean concentration vector $\mu(t)$ with elements

$$\mu_i(t) = \langle X_i(t) \rangle = \frac{\langle N_i(t) \rangle}{\Omega_i}$$

and the concentration covariance matrix $\sigma(t)$ with elements

$$\sigma_{ik}(t) = \langle X_i(t), X_k(t) \rangle = \frac{\langle N_i(t), N_k(t) \rangle}{\Omega_i \Omega_k}$$

When obvious from the context, we will leave out dependence on time.

The two-moment equations (3.7) and (3.14), in the simplified notation, take the form

$$\frac{\mathrm{d}\mu_i}{\mathrm{d}t} = \tilde{f}_i(\mu) + \frac{1}{2} \frac{\partial^2 \tilde{f}_i}{\partial x \partial x^T} : \sigma \tag{4.1}$$

$$\frac{\mathrm{d}\sigma_{ik}}{\mathrm{d}t} = \sum_l \left[\frac{\partial \tilde{f}_i}{\partial x_l} \sigma_{lk} + \sigma_{il} \frac{\partial \tilde{f}_k}{\partial x_l} \right] + \frac{1}{\sqrt{\Omega_i \Omega_k}} \left[\tilde{B}_{ik}(\mu) + \frac{1}{2} \frac{\partial^2 \tilde{B}_{ik}}{\partial x \partial x^T} : \sigma \right] \tag{4.2}$$

where

$$\begin{aligned} \tilde{f}_i(x) &= \frac{1}{\Omega_i} \sum_{j=1}^{r} S_{ij} a_j(\Omega \odot x) \\ \tilde{B}_{ik}(x) &= \frac{1}{\sqrt{\Omega_i \Omega_k}} \sum_{j=1}^{r} S_{ij} S_{kj} a_j(\Omega \odot x) \end{aligned} \tag{4.3}$$

The *effective flux* on the right in (4.1) is the sum of a deterministic flux $\tilde{f}_i(\mu)$ and a stochastic flux $\frac{1}{2} \frac{\partial^2 \tilde{f}_i}{\partial x \partial x^T} : \sigma$, the latter determined by the dynamics of both the mean and (co)variance. This influence of the (co)variance implies that knowledge of fluctuations

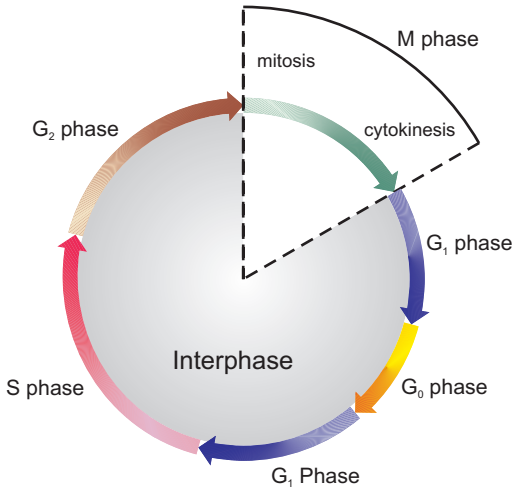

Figure 4.1
Phases of cell cycle regulation. Adopted from [AJL+02].

is important for a correct description of the mean. This also indicates an advantage of the stochastic framework over its deterministic counterpart: starting from the same assumptions and approximations, the stochastic framework allows us to describe the influence of fluctuations on the mean. This can be posed as the central phenomenological argument for stochastic modelling.

The scaling by Ω confirms the inverse relationship between the noise, as measured by (co)variance, and the system size. Note the influence of the mean on the (co)variance in (4.2).

Since the 2MA approach is based on the truncation of terms containing 3rd and higher-order moments, any conclusion from the solution of 2MA must be drawn with care. Ideally, the 2MA should be complemented and checked with a reasonable number of SSA runs.

In [GUV07, Gou07], the 2MA has been applied biochemical systems, demonstrating quantitative and qualitative differences between the mean of the stochastic model and the solution of the deterministic model. The examples used in [GUV07, Gou07] all assume elementary reactions (and hence propensities at most quadratic) and the usual interpretation of concentration as the moles per unit volume. In the next section, we investigate the 2MA for complex systems with non-elementary and relative concentrations. The reason for our interest in non-elementary reactions is the frequent occurrence of rational propensities (reaction rates), e.g. Michaelis-Menten type and Hill type, in models in the system biology literature (e.g. [TCN03]).

4.2 Fission yeast cell cycle modelling

The growth and reproduction of organisms requires a precisely controlled sequence of events known as the cell cycle [AJL+02, Mor07]. On a coarse scale, the cell cycle is composed of four phases: the replication of DNA (S phase), the separation of DNA (mitosis, M phase), and the intervening phases (gaps G1 and G2) which allow for preparation, regulation and control of cell division. These phases are illustrated in Figure 4.1 for a generic cell

cycle. The central molecular components of cell cycle control system have been identified [Nur00, Mor07].

Cell cycle experiments show that cycle times (CTs) have different patterns for the wild type and for various mutants [SNM96, SN02]. For the wild type, the CTs have almost a constant value near 150 min ensured by a size control mechanism: mitosis happens only when the cell has reached a critical size. The double-mutant of fission yeast (namely *wee1⁻ cdc25*Δ) exhibits quantised cycle times: the CTs get clustered into three different groups (with mean CTs of 90, 160 and 230 min). The proposed explanation for the quantised cycle times is a weakend positive feedback loop (due to wee1 and cdc25) which means cells reset (more than once) back to G2 from early stages of mitosis by premature activation of a negative feedback loop [SCNG$^+$00, SN02].

Many deterministic ODE models describing the cell cycle dynamics have been constructed [NCNG$^+$98, NPCT01, NT03, TCNN02]. These models can explain many aspects of the cell cycle including the size control for both the wild type and mutants. Since deterministic models describe the behaviour of a non-existing 'average cell', neglecting the differences among cells in culture, they fail to explain curious behaviours such as the quantised cycle times in the double-mutant. To account for such curiosities in experiments, two stochastic models were constructed by Sveiczer: The first model [SCNG$^+$00, SN02] introduces (external) noise into the rate parameter of the protein Pyp3. The second model [STN01] introduces noise into two cell and nuclear sizes after division asymmetry. Full stochastic models that treat all the time-varying protein concentrations as random variables are reported in [YJT$^+$08, Ste04]. They provide a reasonable explanation for the size control in wild type and the quantised CTs in the double-mutant type. Both models employ the Langevin approach and hence require many simulation runs to provide an ensemble for computing the mean and (co)variance. However, the simulation results of stochastic models in [SCNG$^+$00, SN02, STN01, Ste04, YJT$^+$08] represent one trajectory (for a large number of successive cycles) of the many possible in the ensemble from which the CT statistics (time averages) are computed. We will see that the time-averages computed from the 2MA simulation are for the ensemble of all trajectories.

4.2.1 The deterministic cell cycle model

We base our 2MA model on the deterministic ODE model for the fission yeast cell cycle, developed by Tyson-Novák in [NPCT01]. As shown in Figure 4.2, the cell cycle control mechanism centres around the M-phase promoting factor (MPF), the active form of the heterodimer Cdc13/Cdc2, and its antagonistic interactions with enemies (Ste9, Slp1, Rum1) and the positive feedback with its friend Cdc25. These interactions, among many others, define a sequence of check points to control the timing of cell cycle phases. The result is MPF activity oscillation between low (G1-phase), intermediate (S- and G2-phases) and high (M-phase) levels that is required for the correct sequence of cell cycle events. For simplicity, it is assumed that the cell divides functionally when MPF drops from 0.1.

Table 4.1 lists the proteins whose concentrations x_i, together with MPF concentration, are treated as dynamic variables that evolve according to

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \tilde{f}_i^+(x) - \tilde{f}_i^-(x)\,. \tag{4.4}$$

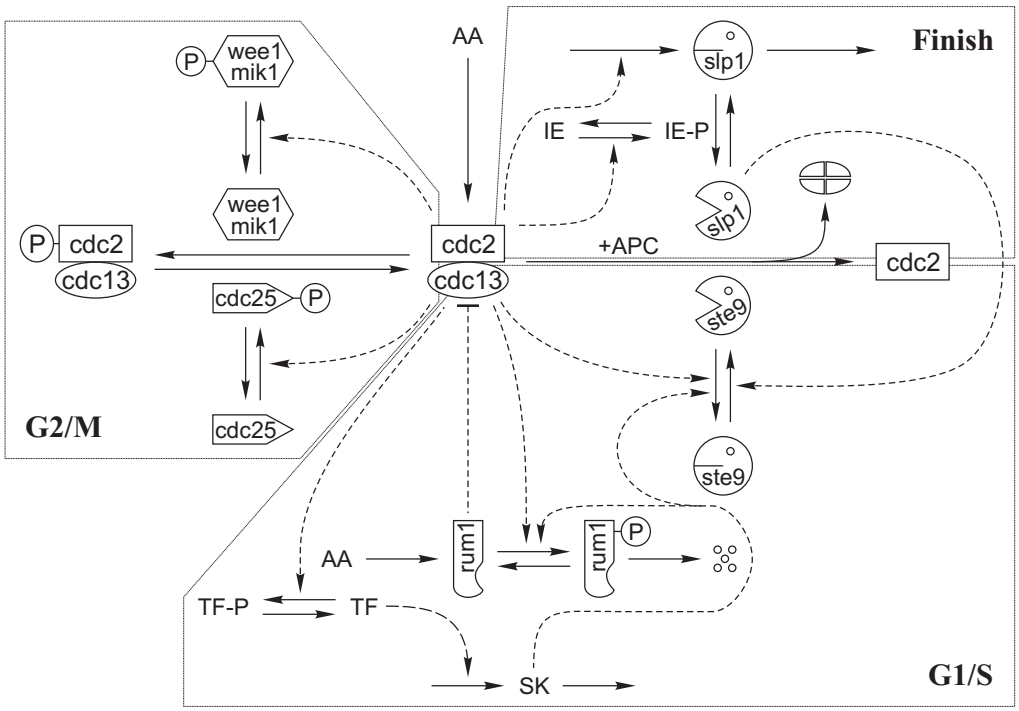

Figure 4.2 Regulation of fission yeast cell cycle. Reproduced from [NPCT01].

Table 4.1 Proteins and fluxes. Here x denotes the vector of concentrations x_1 to x_8.

Index i	Protein	Production flux $\tilde{f}_i^+(x)$	Elimination flux $\tilde{f}_i^-(x)$
1	$\mathrm{Cdc13_T}$	$k_1 M$	$(k_2' + k_2'' x_3 + k_2''' x_5)\, x_1$
2	preMPF	$(x_1 - x_2)\, k_{\mathrm{wee}}$	$(k_{25} + k_2' + k_2'' x_3 + k_2''' x_5)\, x_2$
3	Ste9	$\frac{(k_3' + k_3'' x_5)(1 - x_3)}{J_3 + 1 - x_3}$	$\frac{(k_4' x_8 + k_4 x_{\mathrm{mpf}}) x_3}{J_4 + x_3}$
4	$\mathrm{Slp1_T}$	$k_5' + \frac{k_5'' x_{\mathrm{mpf}}^4}{J_4^4 + x_{\mathrm{mpf}}^4}$	$k_6 x_4$
5	Slp1	$k_7 \frac{(x_4 - x_5) x_6}{J_7 + x_4 - x_5}$	$k_6 x_5 + k_8 \frac{x_5}{J_8 + x_5}$
6	IEP	$k_9 \frac{(1 - x_6) x_{\mathrm{mpf}}}{J_9 + 1 - x_6}$	$k_{10} \frac{x_6}{J_{10} + x_6}$
7	$\mathrm{Rum1_T}$	k_{11}	$(k_{12} + k_{12}' x_8 + k_2'' x_{\mathrm{mpf}})\, x_7$
8	SK	$k_{13} x_{\mathrm{tf}}$	$k_{14} x_8$

Table 4.2 Parameter values for the Tyson-Novák cell cycle model of the fission yeast (wild type) [NPCT01]. All constants have units min^{-1}, except the Js, which are dimensionless Michaelis constants, and K_{diss}, which is a dimensionless equilibrium constant for trimer dissociation. For the double-mutant type, one makes the following three changes: $k''_{\text{wee}} = 0.3$, $k'_{25} = k''_{25} = 0.02$.

$k_{15} = 0.03$, $k'_2 = 0.03$, $k''_2 = 1$, $k'''_2 = 0.1$, $k'_3 = 1$, $k''_3 = 10$, $J_3 = 0.01$,
$k'_4 = 2$, $k_4 = 35$, $J_4 = 0.01$, $k'_5 = 0.005$, $k''_5 = 0.3$, $k_6 = 0.1$, $J_5 = 0.3$,
$k_7 = 1$, $k_8 = 0.25$, $J_7 = J_8 = 0.001$, $J_8 = 0.001$, $k_9 = 0.1$, $k_{10} = 0.04$,
$J_9 = 0.01$, $J_{10} = 0.01$, $k_{11} = 0.1$, $k_{12} = 0.01$, $k'_{12} = 1$, $k''_{12} = 3$, $K_{\text{diss}} = 0.001$,
$k_{13} = 0.1$, $k_{14} = 0.1$, $k_{15} = 1.5$, $k'_{16} = 1$, $k''_{16} = 2$, $J_{15} = 0.01$, $J_{16} = 0.01$,
$V_{\text{awee}} = 0.25$, $V_{\text{iwee}} = 1$, $J_{\text{awee}} = 0.01$, $J_{\text{iwee}} = 0.01$, $V_{\text{a25}} = 1$, $V_{\text{i25}} = 0.25$,
$J_{\text{a25}} = 0.01$, $J_{\text{i25}} = 0.01$, $k'_{\text{wee}} = 0.15$, $k''_{\text{wee}} = 1.3$, $k'_{25} = 0.05$, $k''_{25} = 5$, $\rho = 0.005$

Here $\tilde{f}_i^+(x)$ is the production flux and $\tilde{f}_i^-(x)$ is the elimination flux of ith protein. Note that the summands in the fluxes $\tilde{f}_i^+(x)$ and $\tilde{f}_i^-(x)$ are rates of reactions, most of which, are non-elementary (summarising many elementary reactions into a single step). Quite a few of these reaction rates have rational expressions which requires the extended 2MA approach developed in this paper. The MPF concentration x_{mpf} can be obtained from the algebraic relation

$$x_{\text{mpf}} = \frac{(x_1 - x_2)\,(x_1 - x_{\text{trim}})}{x_1} \tag{4.5}$$

where

$$\begin{aligned}
\frac{\mathrm{d}M}{\mathrm{d}t} &= \rho M \\
x_{\text{trim}} &= \frac{2x_1x_7}{\Sigma + \sqrt{\Sigma^2 - 4x_1x_7}} \\
x_{\text{tf}} &= G\left(k_{15}M, k'_{16}, k''_{16}x_{\text{mpf}}, J_{15}, J_{16}\right) \\
k_{\text{wee}} &= k'_{\text{wee}} + \left(k''_{\text{wee}} - k'_{\text{wee}}\right) G\left(V_{\text{awee}}, V_{\text{iwee}}x_{\text{mpf}}, J_{\text{awee}}, J_{\text{iwee}}\right) \\
k_{25} &= k'_{25} + \left(k''_{25} - k'_{25}\right) G\left(V_{\text{a25}}x_{\text{mpf}}, V_{\text{i25}}, J_{\text{a25}}, J_{\text{i25}}\right) \\
\Sigma &= x_1 + x_7 + K_{\text{diss}}, \\
G(a,b,c,d) &= \frac{2ad}{b - a + bc + ad + \sqrt{(b - a + bc + ad)^2 - 4(b-a)ad}}
\end{aligned} \tag{4.6}$$

Note that the cellular mass M is assumed to grow exponentially with a rate ρ, and the concentrations $(x_{\text{trim}}, x_{\text{tf}}, k_{\text{wee}}, k_{25})$ are assumed to be in a pseudo-steady-state to simplify the model. Note that we use a slightly different notation: ρ for mass growth rate (instead of μ), x_{trim} for Trimmer concentration and x_{tf} for TF concentration. We have to emphasise that the concentrations used in this model are relative and dimensionless. When one concentration is divided by another, the proportion is the same as a proportion of two copy numbers. Hence, such a concentration should not be interpreted as a copy number per unit volume (as misinterpreted in [YJT+08]). The parameters used in the Tyson-Novák model [NPCT01] are listed in Table 4.2.

The deterministic ODE model describes the behaviour of an 'average cell', neglecting the differences among cells in culture. Specifically, it fails to explain the experimentally observed clusters of the CT-vs-BM plot and the tri-modal distribution of CT [SNM96, SCNG$^+$00, STN01, SN02].

4.2.2 Feasibility of Gillespie simulations

Ideally, we should repeat many runs of Gillespie's SSA and compute our desired moments from the ensemble of those runs. At present, there are two problems which this. The first problem is the requirement of elementary reactions for SSA. The elementary reactions underlying the deterministic model [NPCT01] are not known. Many elementary steps have been simplified to obtain that model. Trying to perform SSA on non-elementary reactions is not an option because that will lose the discrete event character of SSA. The second problem arises from the fact that the SSA requires copy numbers which in turn requires knowledge of measured concentrations. All protein concentrations in the model are expressed in arbitrary units (a.u.) because the actual concentrations of most regulatory proteins in the cell are not known [CNBC$^+$06]. Tyson and Sveiczer[1] define relative concentration x_i of the ith protein as $x_i = n_i/\Omega_i$ where $\Omega_i = C_i N_\mathrm{A} V$. Here C_i is an unknown characteristic concentration of the ith component. The idea is to make the relative concentrations x_i free of scale of the actual (molar) concentrations $n_i/N_\mathrm{A}V$. Although one would like to vary C_i, this is computationally intensive. This problem is not so serious for the continuous approximations such as CLE, LNA and the 2MA which are all ODEs and can be numerically solved.

4.2.3 The stochastic model using Langevin's approach

In [YJT$^+$08] a stochastic model is proposed that replaces the ODE model (4.4) with a set of chemical Langevin equations (CLEs)

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \tilde{f}_i^+(x) - \tilde{f}_i^-(x) + \frac{1}{\Omega}\left[\sqrt{\tilde{f}_i^+(x)}\Gamma_i^+(t) - \sqrt{\tilde{f}_i^-(x)}\Gamma_i^-(t)\right],$$

which uses the Langevin noise terms: White noises Γ_i^+ and Γ_i^- scaled by $\sqrt{\tilde{f}_i^+(x)}$ and $\sqrt{\tilde{f}_i^-(x)}$ to represent the internal noise. The system parameter Ω has been described as the volume by the author. As we discussed before, the concentrations are relative levels with different system size parameters. That means that concentrations are not the same as copy numbers per unit volume.

Another stochastic model employing the Langevin's approach is reported in [Ste04] which approximates the squared noise amplitudes by linear functions:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \tilde{f}_i\left(x(t)\right) + \sqrt{2D_i x_i}\Gamma_i(t),$$

where D_i is a constant. The reason why the model dynamics $\tilde{f}(x)$ are missing in this model is that the author wanted to represent both the internal and external noise by the second term on the right.

[1] Personal communication.

Table 4.3 Rows of the drift matrix of the 2MA cell cycle model.

i	$\frac{\partial \tilde{f}_i}{\partial x^T}$
1	$\left[-k_2' - k_2''\mu_3 - k_2'''\mu_5, 0, -k_2''\mu_1, 0, -k_2'''\mu_1, 0, 0, 0\right]$
2	$\left[k_{\text{wee}}, -k_{\text{wee}} - k_{25} - k_2' - k_2''\mu_3 - k_2'''\mu_5, -k_2''\mu_2, 0, -k_2'''\mu_2, 0, 0, 0\right]$
3	$\left[0, 0, -\frac{\left(k_4'\mu_8 + k_4\mu_{\text{mpf}}\right)J_4}{\left(J_4+\mu_3\right)^2} - \frac{\left(k_3'+k_3''\mu_5\right)J_3}{\left(J_3+1-\mu_3\right)^2}, 0, \frac{(1-\mu_3)k_3''}{J_3+1-\mu_3}, 0, 0, -\frac{k_4'\mu_3}{J_4+\mu_3}\right]$
4	$[0, 0, 0, -k_6, 0, 0, 0, 0]$
5	$\left[0, 0, 0, \frac{k_7 J_7 \mu_6}{\left(J_7+\mu_4-\mu_5\right)^2}, -k_6 - \frac{k_7 J_7 \mu_6}{\left(J_7+\mu_4-\mu_5\right)^2} - \frac{k_8 J_8}{\left(J_8+\mu_5\right)^2}, \frac{(\mu_4-\mu_5)k_7}{J_7+x_4-\mu_5}, 0, 0\right]$
6	$\left[0, 0, 0, 0, 0, -\frac{k_9 x_{\text{mpf}} J_9}{\left(J_9+1-\mu_6\right)^2} - \frac{k_{10} J_{10}}{\left(J_{10}+\mu_6\right)^2}, 0, 0\right]$
7	$\left[0, 0, 0, 0, 0, 0, -k_{12} - k_{12}'\mu_8 - k_2''\mu_{\text{mpf}}, -k_{12}'\mu_7\right]$
8	$\left[0, 0, 0, 0, 0, 0, 0, -k_{14}\right]$

4.2.4 The 2MA cell cycle model

For the cell cycle model, the flux $\tilde{f}$ and the diffusion matrix $\tilde{B}$, defined in (4.3), have elements

$$\tilde{f}_i(x) = \tilde{f}_i^+(x) - \tilde{f}_i^-(x), \quad \tilde{B}_{ik}(x) = \begin{cases} \tilde{f}_i^+(x) + \tilde{f}_i^-(x) & \text{if } i = k \\ 0 & \text{if } i \neq k\,. \end{cases}$$

The off-diagonal elements of $\tilde{B}$ are zero because each reaction changes only one component, so that $S_{ij}S_{kj} = 0$ for $i \neq k$. Once these quantities are known, it follows from (4.1) and (4.2) that the following set of ODEs:

$$\frac{\mathrm{d}\mu_i}{\mathrm{d}t} = \tilde{f}_i(\mu) + \frac{1}{2}\frac{\partial^2 \tilde{f}_i}{\partial x \partial x^T} : \sigma \tag{4.7}$$

$$\frac{\mathrm{d}\sigma_{ii}}{\mathrm{d}t} = 2\sum_l \frac{\partial \tilde{f}_i}{\partial x_l}\sigma_{li} + \frac{1}{\Omega_i}\left[\tilde{B}_{ii}(\mu) + \frac{1}{2}\frac{\partial^2 \tilde{B}_{ii}}{\partial x \partial x^T} : \sigma\right] \tag{4.8}$$

$$\frac{\mathrm{d}\sigma_{ik}}{\mathrm{d}t} = \sum_l \left[\frac{\partial \tilde{f}_i}{\partial x_l}\sigma_{lk} + \sigma_{il}\frac{\partial \tilde{f}_k}{\partial x_l}\right] \quad i \neq k \tag{4.9}$$

approximates (correctly to the 2nd-order moments) the evolution of component-wise concentration mean and covariance. See Table 4.3 for the respective expressions of the drift matrix, Table 4.4 the stochastic flux and Table 4.5 for the 2nd-order term in the Taylor expansion of $\tilde{B}_{ii}$ in (4.8).

Table 4.4 Stochastic flux, the 2nd-order term in the Taylor expansion of $\tilde{f}_i$ around the mean.

i	$\frac{1}{2}\frac{\partial^2 f_i}{\partial x \partial x^T} : \sigma$
1	$-k_2''\sigma_{13} - k_2'''\sigma_{15}$
2	$-k_2''\sigma_{23} - k_2'''\sigma_{25}$
3	$\left[\frac{\left(k_4'\mu_8+k_4\mu_{\mathrm{mpf}}\right)J_4}{\left(J_4+\mu_3\right)^3} - \frac{\left(k_3'+k_3''\mu_5\right)J_3}{\left(J_3+1-\mu_3\right)^3}\right]\sigma_{33} - \frac{k_3''J_3\sigma_{35}}{\left(J_3+1-\mu_3\right)^2} - \frac{k_4'J_4\sigma_{38}}{\left(J_4+\mu_3\right)^2}$
4	0
5	$\frac{k_7J_7\mu_6(2\sigma_{45}-\sigma_{44}-\sigma_{55})}{\left(J_7+\mu_4-\mu_5\right)^3} + \frac{k_7J_7(\sigma_{46}-\sigma_{56})}{\left(J_7+\mu_4-\mu_5\right)^2} + \frac{k_8J_8}{\left(J_8+\mu_5\right)^3}\sigma_{55}$
6	$\left[\frac{k_{10}J_{10}}{\left(J_{10}+\mu_6\right)^3} - \frac{k_9\mu_{\mathrm{mpf}}J_9}{\left(J_9+1-\mu_6\right)^3}\right]\sigma_{66}$
7	$-k_{12}'\sigma_{78}$
8	0

Table 4.5 The 2nd-order term in the Taylor expansion of $\tilde{B}_{ii}$ around the mean.

i	$\frac{1}{2}\frac{\partial^2 \tilde{B}_{ii}}{\partial x \partial x^T} : \sigma$
1	$k_2''\sigma_{13} + k_2'''\sigma_{15}$
2	$k_2''\sigma_{23} + k_2'''\sigma_{25}$
3	$-\left[\frac{\left(k_4'\mu_8+k_4\mu_{\mathrm{mpf}}\right)J_4}{\left(J_4+\mu_3\right)^3} + \frac{\left(k_3'+k_3''\mu_5\right)J_3}{\left(J_3+1-\mu_3\right)^3}\right]\sigma_{33} - \frac{k_3''J_3\sigma_{35}}{\left(J_3+1-\mu_3\right)^2} + \frac{k_4'J_4\sigma_{38}}{\left(J_4+\mu_3\right)^2}$
4	0
5	$\frac{k_7J_7\mu_6(2\sigma_{45}-\sigma_{44}-\sigma_{55})}{\left(J_7+\mu_4-\mu_5\right)^3} + \frac{k_7J_7(\sigma_{46}-\sigma_{56})}{\left(J_7+\mu_4-\mu_5\right)^2} - \frac{k_8J_8}{\left(J_8+\mu_5\right)^3}\sigma_{55}$
6	$-\left[\frac{k_{10}J_{10}}{\left(J_{10}+\mu_6\right)^3} + \frac{k_9\mu_{\mathrm{mpf}}J_9}{\left(J_9+1-\mu_6\right)^3}\right]\sigma_{66}$
7	$k_{12}'\sigma_{78}$
8	0

Having at hand the moments involving the eight dynamic variables x_1 to x_8, the mean MPF concentration can also be approximated. Towards that end, we start with the MPF concentration

$$x_{\mathrm{mpf}} = (x_1 - x_2)\left(1 - \frac{x_{\mathrm{trim}}}{x_1}\right) = x_1 - x_2 - x_{\mathrm{trim}} + x_{\mathrm{trim}}\frac{x_2}{x_1}\,.$$

The ratio x_2/x_1 can be expanded around the mean,

$$\frac{x_2}{x_1} = \frac{1}{\mu_1}\frac{x_2}{1+\frac{(x_1-\mu_1)}{\mu_1}} = \frac{1}{\mu_1}\left[x_2 - \frac{(x_1-\mu_1)\,x_2}{\mu_1} + \frac{(x_1-\mu_1)^2\,x_2}{\mu_1^2} + \cdots\right].$$

Taking expectation on both sides,

$$\begin{aligned}\left\langle\frac{X_2}{X_1}\right\rangle &= \frac{1}{\mu_1}\left\langle\frac{X_2}{1+\frac{(X_1-\mu_1)}{\mu_1}}\right\rangle \\ &= \frac{1}{\mu_1}\left\langle X_2 - \frac{(X_1-\mu_1)\,X_2}{\mu_1} + \frac{(X_1-\mu_1)^2\,X_2}{\mu_1^2} + \cdots\right\rangle \\ &= \frac{1}{\mu_1}\left[\mu_2 - \frac{\sigma_{12}}{\mu_1} + \frac{\mu_2\sigma_{11}}{\mu_1^2}\right].\end{aligned}$$

Finally, with the understanding that x_{trim} is in pseudo steady state, the mean MPF concentration follows from the expectation of x_{mpf} to be

$$\mu_{\mathrm{mpf}} = \mu_1 - \mu_2 - x_{\mathrm{trim}} + \frac{x_{\mathrm{trim}}}{\mu_1}\left[\left(1+\frac{\sigma_{11}}{\mu_1^2}\right)\mu_2 - \frac{\sigma_{12}}{\mu_1}\right]. \tag{4.10}$$

This expression for the average MPF activity demonstrates the influence of (co)variance on the mean as emphasised here. We see the dependence of mean MPF concentration μ_{mpf} on the variance σ_{11} and covariance σ_{12} in addition to the means μ_1, μ_2 and x_{trim}.

4.2.5 Simulations of the 2MA model

The system of ODEs (4.7)-(4.9) was solved numerically by the MATLAB solver `ode15s` [Mat07]. The solution was then combined with algebraic relations (4.10). For parameter values, see Table 4.2. Since information about the individual scaling parameters Ω_i used in the definition of concentrations is not available, we have used $\Omega_i = 5000$ for all i. Note, however, that the 2MA approach developed here will work for any combination of $\{\Omega_i\}$. The time courses of mass and MPF activity are plotted in Figure 4.3a for the wild type and in Figure 4.3b for the double-mutant type. For the wild type, the 2MA predicted mean trajectories do not differ considerably from the corresponding deterministic trajectories. Both show a constant CT of near 150 min. Thus internal noise does not seem to have a major influence for the wild type.

For the double-mutant type, the difference between the 2MA and deterministic predictions is significant. The deterministic model (4.4) predicts alternating short cycles and long cycles because cells born at the larger size have shorter cycle, and smaller newborns have longer cycles [NPCT01]. This strict alternation due to size control is not observed

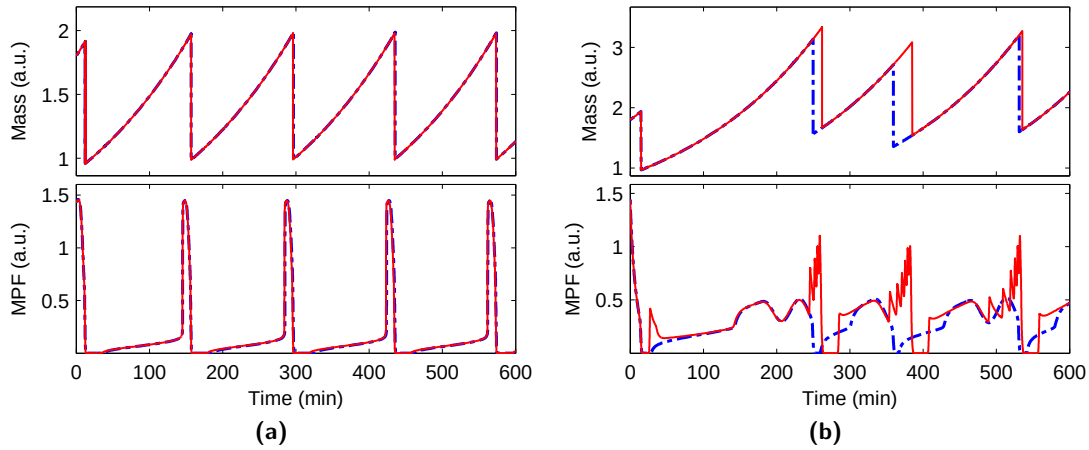

Figure 4.3 The time courses of mass and MPF activity: (a) for the wild type, (b) for the double-mutant type. The 2MA predicted mean trajectories are plotted as solid lines and the corresponding deterministic trajectories as dashed lines. The difference between the two predictions is negligible for the wild type, but significant for double-mutant type.

in experiments: cells of same mass may have short or long cycles (excluding very large cells that have always the shortest CT) [SNM96, SCNG$^+$00]. This lack of size control is reproduced by the 2MA simulations: the multiple resettings of MPF to G2, induced by the internal noise, result in longer CTs (thus accounting for the 230 min cycles observed experimentally). Such MPF resettings have been proposed in [SCNG$^+$00, SN02] to explain quantised CTs. No such resetting is demonstrated by the deterministic model.

Figure 4.4 additionally shows time courses of slp1, Ste9 and Rum1$_\mathrm{T}$. For the wild-type, the difference in the Rum1T concentrations near the G2/M transition has no significant effect on the MPF activity because Rum1T tries to inhibit MPF in G2-phase. For the double-mutant type, the oscillatory behaviour of Ste9 and Slp1 may have resulted in the oscillatory behaviour of the MPF near the G2/M transition which in turn delays the mitosis by a noticeable period.

Note that the mean $\mu(t)$ of the 2MA describes the average of an ensemble of cells. Yet the MPF resettings observed in Figure (4.3b), near G2/M transition, introduce the required variability that explains the clustering of the cycle time observed in experiments. This is in contrast to the alternative stochastic approaches in [SCNG$^+$00, SN02, STN01, Ste04, YJT$^+$08] that use one sample trajectory rather than the ensemble average.

How do we explain this significant effect of noise for the double-mutant on one hand and its negligible effect for the wild type on the other hand? If we look at expression (4.10), we see the influence of the variance σ_{11} (of Cdc13$_\mathrm{T}$) and covariance σ_{12} (between Cdc13$_\mathrm{T}$ and preMPF) on the mean MPF concentration μ_{mpf}. The two (co)variances are plotted in Figure 4.5a for the wild type and in Figure 4.5b for the double-mutant type. It is clear that the two (co)variances have very small peaks for the wild type compared to the large peaks for the double-mutant type. Note that the larger peaks in Figure 4.5b are located at the same time points where the MPF activity exhibits oscillations and hence multiple resettings to G2. This suggest that the oscillatory behaviour of MPF near the

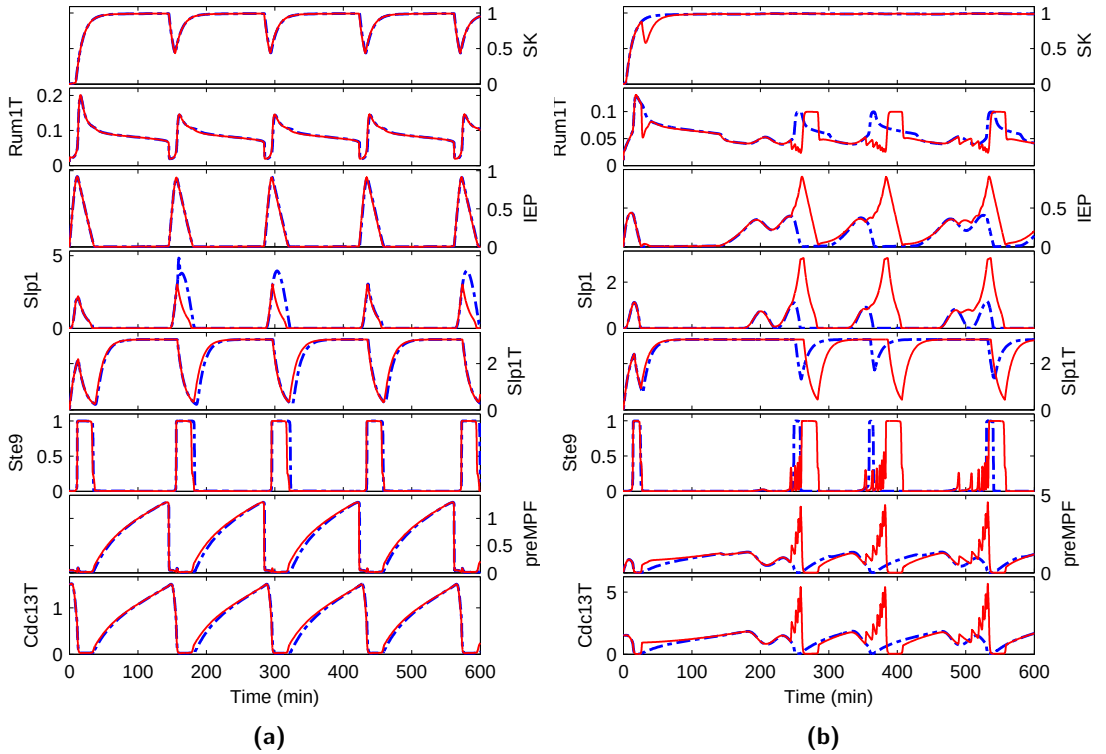

Figure 4.4 The time courses of protein concentrations: (a) for the wild type, (b) for the double-mutant type. The 2MA predicted mean trajectories are plotted as solid lines and the corresponding deterministic trajectories as dashed lines.

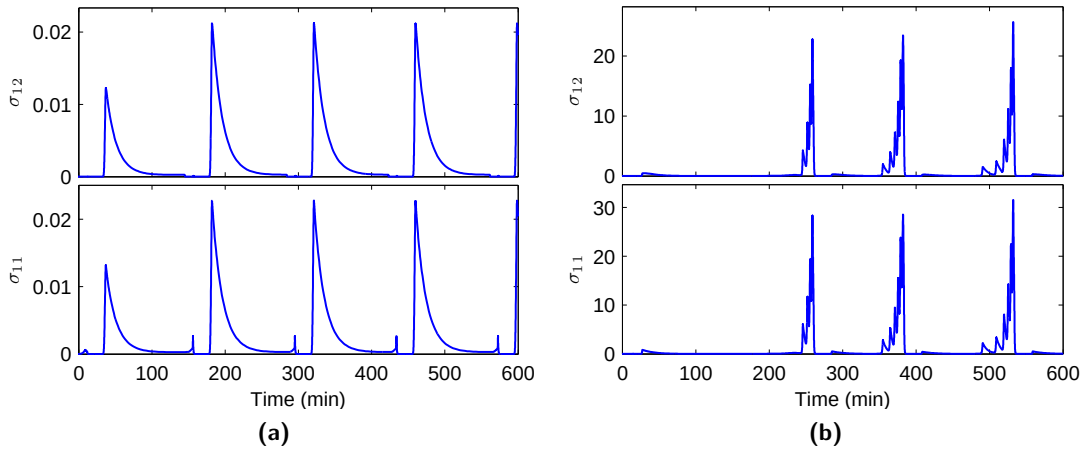

Figure 4.5 Variance σ_{11} (of $\mathrm{Cdc13_T}$) and covariance σ_{12} (between $\mathrm{Cdc13_T}$ and preMPF): (a) for the wild type, (b) for double-mutant type.

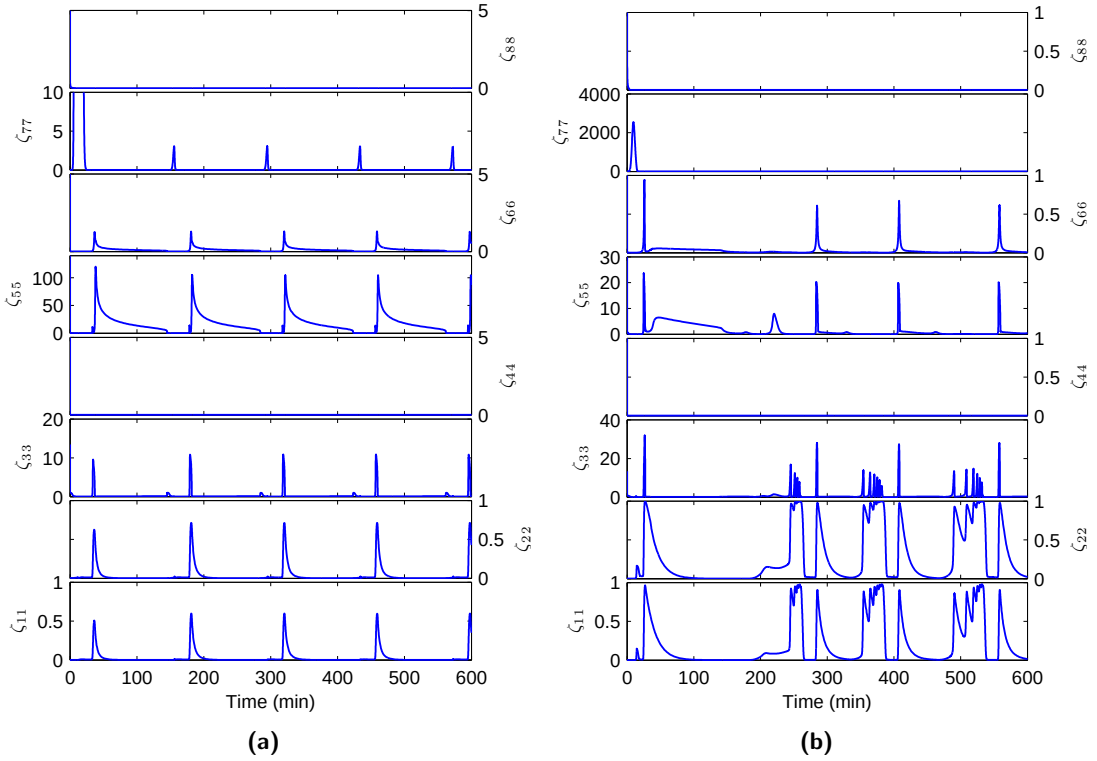

Figure 4.6 Noise-to-signal ratio (NSR): (a) for the wild type, (b) for double-mutant type.

G2/M transition is due to the influence of the oscillatory (co)variances. This coupling between the mean and (co)variance is not captured by the deterministic model.

To allow for comparison between component-wise variances, the variance is usually normalised by the squared mean gives a dimensionless ratio which also removes the dependence of the variance on the scale of the mean. The normalised variance,

$$\zeta_{ii} = \frac{\sigma_{ii}}{\mu_i^2}$$

is known as the signal-to-noise ratio (NSR). The NSR as a measure of noise is usually preferred because of being dimensionless and allowing for additivity of noise and using indirectly measured concentrations [GUV07]. See [Pau05] for different measures of noise and their merits.

The component-wise NSR is plotted in Figure 4.6. We note that the NSR for the double-mutant have irregular oscillations compared to the almost periodic for the wild-type. This may be one of the reasons behind the significant difference in the evolution mean compared to the deterministic evolution for the double-mutant type. The pair-wise variation between components is better described by the cross signal-to-noise ratio (xNSR)

$$\zeta_{ik} = \frac{\sigma_{ik}}{\mu_i^2}\,.$$

The xNSR, for selected component pairs appearing in the expression for stochastic flux (see Table 4.4), is plotted in Figure 4.7. The oscillatory behaviour in both plots suggests

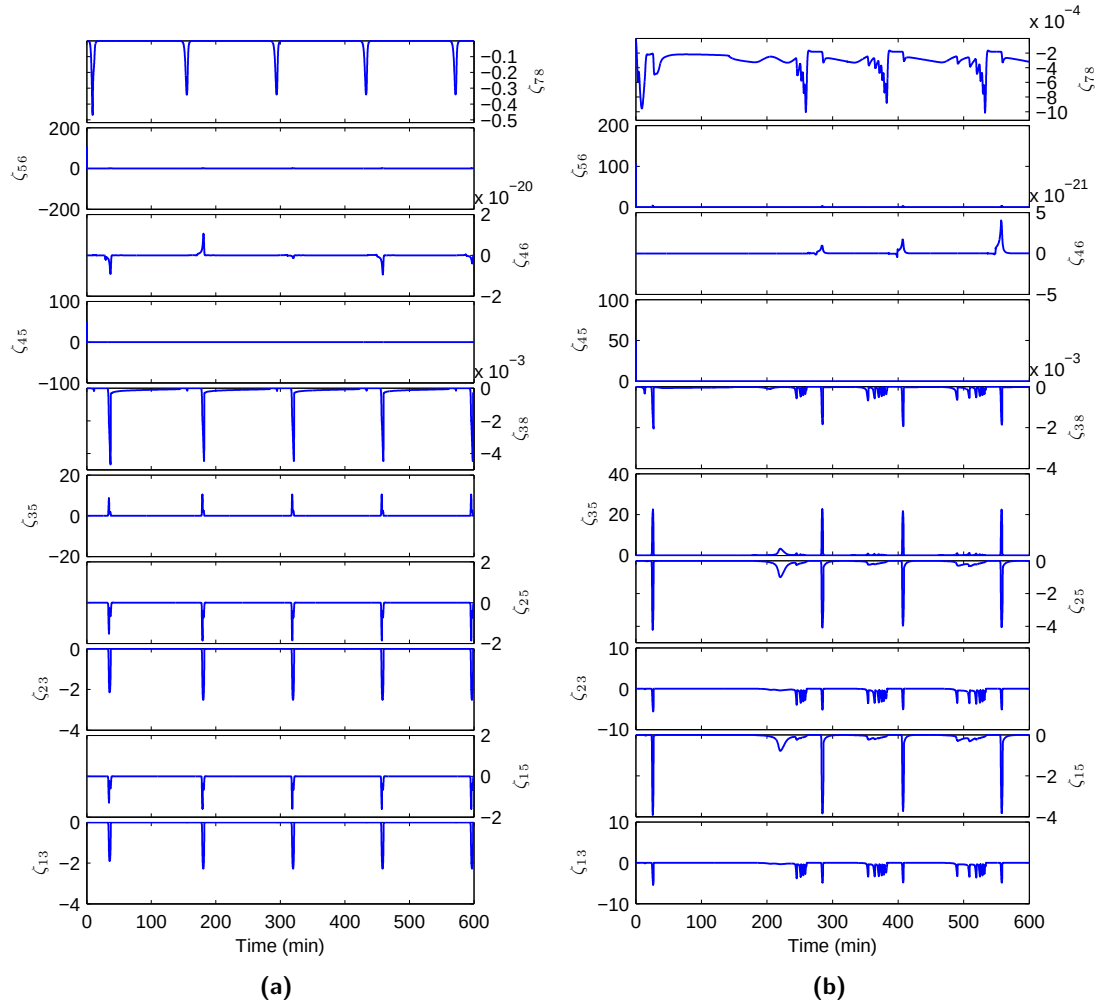

Figure 4.7 Cross noise-to-signal ratio (xNSR) of selected component pairs: (a) for the wild type, (b) for double-mutant type.

Table 4.6 Statistics over 465 successive cell cycles of the double-mutant type cells, predicted by the 2MA model, compared with experimental data, see [SNM96, Table 1].

Case	μ_{CT}	σ_{CT}	$\mathrm{CV}_{\mathrm{CT}}$	μ_{dM}	σ_{dM}	$\mathrm{CV}_{\mathrm{dM}}$	μ_{BM}	σ_{BM}
(1)	131	47	0.358	2.22	0.45	0.203	1.21	0.24
(2)	138.8	12.4	0.09	1.59	0.058	0.0362	3.18	0.101
(3)	138.8	17.6	0.127	1.62	0.093	0.0576	3.25	0.178
(4)	138.8	23.9	0.172	1.66	0.12	0.0721	3.32	0.231

(1) experimental data, (2) $\Omega = 5000$, (3) $\Omega = 5200$, (4) $\Omega = 5300$.

that the off-diagonal elements of the covariance matrix may have influenced the mean on one hand and have a mutual influence on each other. At these points the system is sensitive to noise. Capturing these phenomena is of particular importance if one considers cells in their context (e.g. tissue) where cell-cell variations form the basis for functional mechanisms at higher levels of cellular organisation.

It has to be realised that the above proposition requires validation since the 2MA approach ignores 3rd and higher-order moments. We cannot know whether that truncation is responsible for the oscillations in Figures 4.3 and 4.5, unless compared with a few sample trajectories simulated by the SSA. However, as discussed before, the SSA cannot be performed (at present) for the model in consideration. Therefore we need to compare the 2MA predictions for the double-mutant type cells with experimental data. Towards that end, values of cycle time (CT), birth mass (BM) and division mass (DM) were computed for 465 successive cycles of double-mutant cells. Figure 4.8 shows the CT-vs-BM plot and the CT distribution for three different values $\{5000, 5200, 5300\}$ of system size Ω.

To make this figure comparable with experimental data from [SNM96, SN02], we assume that 1 unit of mass corresponds to 8.2 μm cell length [SCNG$^+$00]. We can see the missing size control (CT clusters), in qualitative agreement with experimentally observed ones (see [SNM96, Figure 6] and [SN02, Figure 5] for a comparison). There are more than four clusters, which may have arisen from the truncated higher-order moments. The extreme value of CT higher than 230 min suggests more than two MPF resettings. Furthermore, more than three modes in the CT distribution may have arisen from the truncated higher-order moments. Table 4.6 compares the statistics for the double-mutant type cells, computed with the 2MA approach, with data from [SNM96, Table 1]. Column 2-4 tabulate, for CT, the mean μ_{CT}, the standard deviation σ_{CT} and the coefficient of variation $\mathrm{CV}_{\mathrm{CT}}$, respectively. The other columns tabulate similar quantities for the division mass (DM) and birth mass (BM). We see that only the mean CT is in agreement with the experimental data. The mean BM is much larger than the experimental BM. The other statistics are much smaller the corresponding experimental values. This table and the above plots suggest that the 2MA should be used with caution. However, another aspect of the cell cycle model deserves attention here. The way the

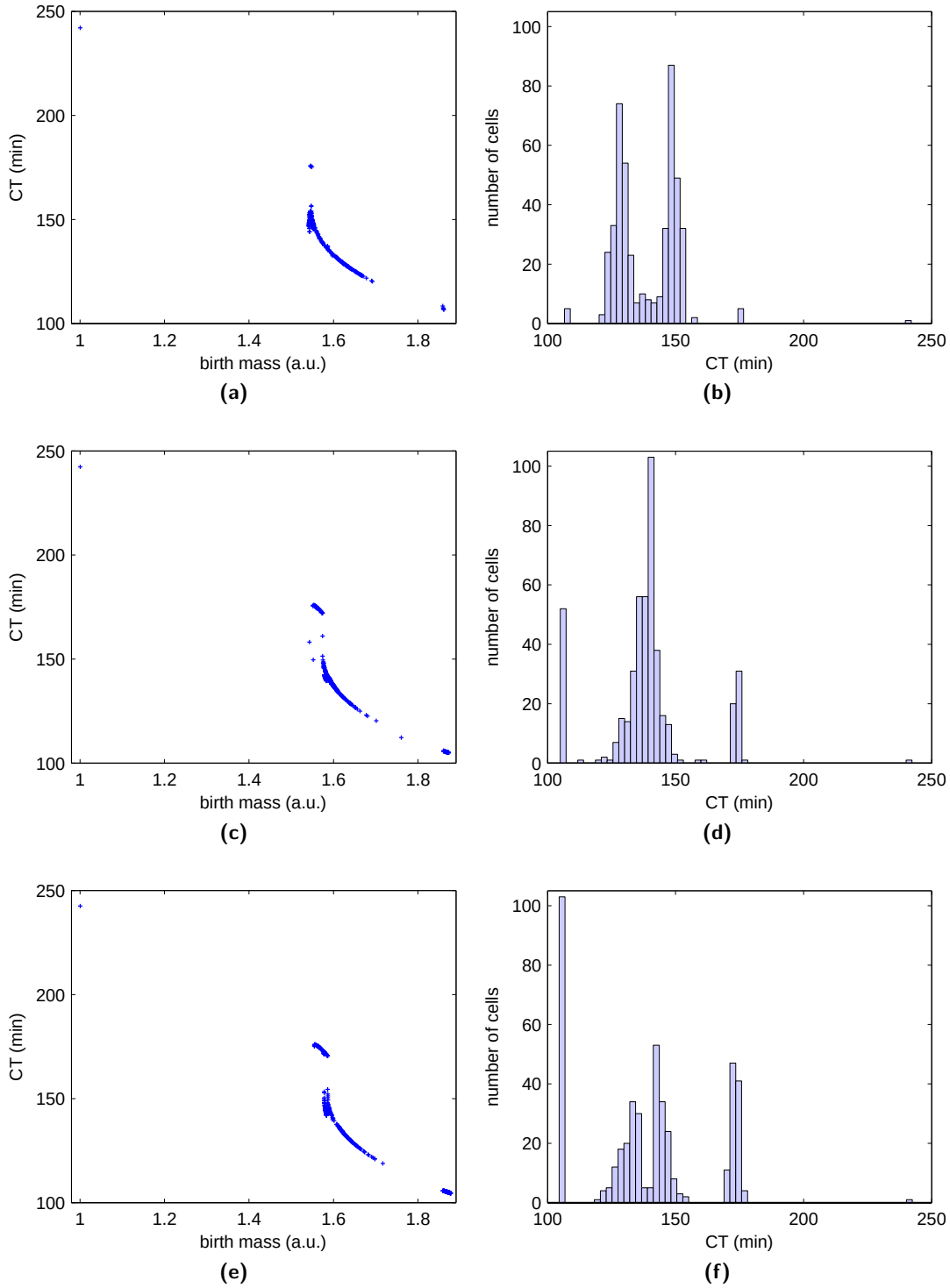

Figure 4.8 Cycle time behaviour over 465 successive cycles of the double-mutant cells, predicted by the 2MA model. (a,c,e): CT vs BM, (b,d,f): CT distribution, (a,b): $\Omega = 5000$, (c,d): $\Omega = 5200$, (e,f): $\Omega = 5300$. The plots are in qualitative agreement to experiments, see Figure 6 in [SNM96] and Figure 5 in [SN02] for a comparison.

relative protein concentrations have been defined implies unknown values of the scaling parameters $\{\Omega_i\}$. Since $\Omega_i = C_i N_A V$, knowing the volume V does not solve the problem: the characteristic concentrations $\{C_i\}$ are still unknown. Our simulations have chosen typical values $\Omega = \{5000, 5200, 5300\}$. The corresponding three pairs of plots in Figure 4.8 and rows in Table 4.6 demonstrate a dependence of the results on a suitable system size. There is no way to confirm these values. The scaling parameters could be regulated in a wider range in order to improve the accuracy of our simulation, motivating future work for us. The conclusion is that the quantitative disagreement of the 2MA predictions can be attributed to two factors: 1) the truncated higher-order moments during the derivation of the 2MA, and (2) the unknown values of scaling parameters.

4.3 Conclusions on the cell cycle model

The recently developed two-moment approximation (2MA) [Gou07, GUV07] is a promising approach because it accounts for the coupling between the means and (co)variances. We have extended the derivation of the 2MA to biochemical networks and established two advances to previous efforts: a) relative concentrations and b) non-elementary reactions. Both aspects are important in systems biology where one is often forced to aggregate elementary reactions into single step reactions. In these situations one cannot assume knowledge of elementary reactions to formulate a stochastic model. Previous derivations assumed elementary reactions and absolute concentrations. However, numerous existing models in systems biology use relative concentrations.

We investigated the applicability of the 2MA approach to the well established fission yeast cell cycle model. The simulations of the 2MA model show oscillatory behaviour near the G2/M transition, which is significantly different from the simulations of deterministic ODE model. One notable aspect of our analytical model is that, although it describes the average of an ensemble, it reproduces enough variability among cycles to reproduce the curious quantised cycle times observed in experiments on double mutants.

5

Conclusions

The discrete and random occurrence far from thermodynamic equilibrium of chemical reactions, and low copy numbers of chemical species, in single cells necessitate stochastic approaches for modelling.

This work presents a stochastic framework for modelling subcellular biochemical systems. In Chapter 2, we make an effort to show how the notion of propensity, the chemical master equation and the stochastic simulation algorithm arise as consequences of the Markov property. This connection is not obvious from the relevant literature in systems biology. Moreover, we sketch the formal relationships between various stochastic approaches referred to in the systems biology literature. Throughout the text we use four simple systems to illustrate ideas and motivate stochastic modelling.

The central theme of the present work, however, is the two-moment approximation (2MA) as a bridge between deterministic and stochastic approaches. The 2MA combines the intuition of derministic models with the represention of noise and variability. In contrast to other stochastic approaches, an analytical 2MA model allows us to study the coupling of mean and co-variance. In Chapter 3, we introduce the 2MA approach and develop extensions to allow (a) non-elementary reactions and (b) relative concentrations.

Both aspects, (a) and (b), are characteristic of the Tyson-Novák model for the fission yeast cell cycle, which we use as a case study in Chapter 4. Our analytical model reproduces the clustering of cycle times observed in experiments. This is explained through multiple resettings of a protein called MPF, caused by the coupling between mean and (co)variance, near the G2/M transition. The conclusion is that the 2MA approach can infer new properties in a single simulation run, something that is neither possible with the deterministic approach (no representation of noise) nor other stochastic approaches (requiring many simulation runs).

With regard to further work, the development of a tool to automate the construction of 2MA equations would be desirable, since the 2MA equations involve complicated symbolic operations. Furthermore, the truncation of higher-order moments used by the 2MA approach suggests the development of some form of error control. Finally, a sensitivity analysis of the mean and (co)variance in terms of the system size would be of interest.

Bibliography

[ABRB+08] Juliet Ansel, Hélène Bottin, Camilo Rodriguez-Beltran, Christelle Damon, Muniyandi Nagarajan, Steffen Fehrmann, Jean François, and Gaël Yvert. Cell-to-Cell Stochastic Variation in Gene Expression Is a Complex Genetic Trait. *PLoS Genet.*, 4(4):e1000049, Apr 2008. doi:10.1371/journal.pgen.1000049. 1.1

[ADKC07] Maxim N. Artyomov, Jayajit Das, Mehran Kardar, and Arup K. Chakraborty. Purely stochastic binary decisions in cell signaling models without underlying deterministic bistabilities. *Proc. Natl. Acad. Sci. U. S. A.*, 104(48):18958–18963, November 2007. doi:10.1073/pnas.0706110104. 1.1

[AJL+02] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walker. *Molecular Biology of the Cell.* Garland Publishers, New York, 4 edition, 2002. 4.1, 4.2

[AMP07] Sébastien Aumaître, Kirone Mallick, and François Pétrélis. Noise-induced bifurcations, multiscaling and on–off intermittency. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07016, 2007. doi:10.1088/1742-5468/2007/07/P07016. 1.1

[BKvO05] Attila Becskei, Benjamin B. Kaufmann, and Alexander van Oudenaarden. Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.*, 37(9):937–944, September 2005. doi:10.1038/ng1616. 1.1, 1.1, 1.2

[Blo06] C. Blomberg. Fluctuations for good and bad: The role of noise in living systems. *Physics of Life Reviews*, 3:133–161, 2006. doi:10.1016/j.plrev.2006.06.001. 1.1

[BPE00] Otto G. Berg, Johan Paulsson, and Måns Ehrenberg. Fluctuations and Quality of Control in Biological Cells: Zero-Order Ultrasensitivity Reinvestigated. *Biophys. J.*, 79(3):1228–1236, 2000. doi:10.1016/S0006-3495(00)76377-6. 1.1

[CB04] Athel Cornish-Bowden. *Fundamentals of Enzyme Kinetics.* Portland Press, third edition, 2004. 2.2

[CNBC+06] Attila Csikász-Nagy, Dorjsuren Battogtokh, Katherine C. Chen, Bela Novák, and John J. Tyson. Analysis of a generic model of eukaryotic cell cycle regulation. *Biophys. J.*, page biophysj.106.081240, 2006. doi:10.1529/biophysj.106.081240. 1.2, 4.2.2

[CW06] B. S. Chen and Y. C. Wang. On the attenuation and amplification of molecular noise in genetic regulatory networks. *BMC Bioinformatics*, 7:52, February 2006. doi:10.1186/1471-2105-7-52. 1.1

[DL93] Marileen Dogterom and Stanislas Leibler. Physical aspects of the growth and regulation of microtubule structures. *Phys. Rev. Lett.*, 70(9):1347, March 1993. doi:10.1103/PhysRevLett.70.1347. 1.1

[DMK+06] Y. Dublanche, K. Michalodimitrakis, N. Kümmerer, M. Foglierini, and L. Serrano. Noise in transcription negative feedback loops: simulation and experimental analysis. *Molecular Systems Biology*, 2:41, 2006. doi:10.1038/msb4100081. 1.1, 1.2

[EE03] Johan Elf and Måns Ehrenberg. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.*, 13(11):2475–2484, November 2003. doi:10.1101/gr.1196503. 1.1, 1.2

[Elf04] Johan Elf. *Intracellular Flows and Fluctuations*. PhD thesis, Uppsala University, Teknisk-naturvetenskapliga vetenskapsområdet, Biology, Department of Cell and Molecular Biology, September 2004. Available from: http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4291. 2.4

[EPBE03] Johan Elf, Johan Paulsson, Otto G. Berg, and Måns Ehrenberg. Near-critical phenomena in intracellular metabolite pools. *Biophys. J.*, 84(1):154–170, January 2003. doi:10.1016/S0006-3495(03)74839-5. 1.1

[Fel97] D. Fell. *Understanding the Control of Metabolism*. Portland Press, 1997. 2.2

[FHG+04] Hunter B. Fraser, Aaron E. Hirsh, Guri Giaever, Jochen Kumm, and Michael B. Eisen. Noise Minimization in Eukaryotic Gene Expression. *PLoS Biol.*, 2(6):e137, June 2004. doi:10.1371/journal.pbio.0020137. 1.1

[FLH07] Lars Ferm, Per Lötstedt, and Andreas Hellander. A Hierarchy of Approximations of the Master Equation Scaled by a Size Parameter. Technical Report 2007-011, Uppsala University, Department of Information Technology, April 2007. Available from: http://www.it.uu.se/research/publications/reports/2007-011/. 1.2

[FX01] James E. E. Ferrell and Wen Xiong. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(1):227–236, March 2001. doi:10.1063/1.1349894. 1.1

[Gar04] C.W. Gardiner. *Handbook of Stochastic Models*. Springer, third edition, 2004. 2.7

[Gil77] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977. doi:10.1021/jp993732q. 1.2, 2.5

[Gil92] D.T. Gillespie. *Markov Processes*. Academic Press, 1992. 2.1

[Gil96] D.T. Gillespie. The multivariate Langevin and Fokker-Planck equations. *American Journal of Physics*, 64(10):1246–1257, 1996. doi:10.1119/1.18387. 2.6, 2.6.3

[Gil00] D.T. Gillespie. The chemical Langevin equation. *J. Chem. Phys.*, 113(1):297–306, 2000. doi:10.1063/1.481811. 1.2, 2.6.3, 2.6.3

[Gou06] John Goutsias. A hidden Markov model for transcriptional regulation in single cells. *IEEE/ACM Trans Comput Biol Bioinform*, 3(1):57–71, 2006. doi:10.1109/TCBB.2006.2. 1.2, 2.6.3

[Gou07] John Goutsias. Classical versus stochastic kinetics modeling of biochemical reaction systems. *Biophys. J.*, 92(7):2350–2365, April 2007. doi:10.1529/biophysj.106.093781. 1.2, 4.1, 4.3

[GUV07] Carlos A. Gómez-Uribe and George C. Verghese. Mass fluctuation kinetics: capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *J. Chem. Phys.*, 126(2):024109, January 2007. doi:10.1063/1.2408422. 1.1, 1.2, 4.1, 4.2.5, 4.3

[HB08] Gil Hornung and Naama Barkai. Noise Propagation and Signaling Sensitivity in Biological Networks: A Role for Positive Feedback. *PLoS Comput. Biol.*, 4(1):e8, Jan 2008. doi:10.1371/journal.pcbi.0040008. 1.1

[Hig01] Desmond J. Higham. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*, 43(3):525–546, 2001. doi:10.1137/S0036144500378302. 2.6.3

[HJ04] F. Hayot and C. Jayaprakash. The linear noise approximation for molecular fluctuations within cells. *Physical Biology*, 1(4):205–210, December 2004. doi:10.1088/1478-3967/1/4/002. 1.2, 2.6

[Hou07] James House. *Principles of Chemical Kinetics*. Academic Press, 2007. 2.2

[HPDC00] Jeff Hasty, Joel Pradines, Milos Dolnik, and J. J. Collins. Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. U. S. A.*, 97(5):2075–2080, 2000. doi:10.1073/pnas.040411297. 1.1

[HS96] R. Heinrich and S. Schuster. *The Regulation of Cellular Systems*. Chapman and Hall, New York, 1996. 2.2

[Kam07a] N.G. van Kampen. *Stochastic Processes in Physics and Chemistry (Third Edition)*. Elsevier Amsterdam, Amsterdam, 2007. Available from: http://www.sciencedirect.com/science/book/9780444529657. 1.2, 2.6.2, 2.6.3

[Kam07b] N.G. van Kampen. The Langevin approach. In *Stochastic Processes in Physics and Chemistry (Third Edition)*, pages 219–243. Elsevier, Amsterdam, 2007. doi:10.1016/B978-044452965-7/50012-X. 1.2

[KEBC05] Mads Kaern, Timothy C. Elston, William J. Blake, and James J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, 6(6):451–464, June 2005. doi:10.1038/nrg1615. 1.1, 1.1, 1.2

[Kei87] Joel Keizer. *Statistical thermodynamics of nonequilibrium processes*. Springer, Berlin, 1987. 3.3

[Lax60] M. Lax. Fluctuations from the nonequilibrium steady state. *Reviews of Modern Physics*, 32(1):25–64, 1960. doi:10.1103/RevModPhys.32.25. 3.3

[LKM07] Joseph Levine, Hao Yuan Kueh, and Leonid Mirny. Intrinsic fluctuations, robustness, and tunability in signaling cycles. *Biophys. J.*, 92(12):4473–4481, 2007. doi:10.1529/biophysj.106.088856. 1.1

[LL08] Qianshu Li and Xiufeng Lang. Internal Noise-Sustained Circadian Rhythms in a Drosophila Model. *Biophys. J.*, 94(6):1983–1994, 2008. arXiv: http://www.biophysj.org/cgi/reprint/94/6/1983.pdf, doi:10.1529/biophysj.107.109611. 1.1

[Lot20] Alfred J. Lotka. Undamped oscillations derived from the law of mass action. *Journal of the American Chemical Society*, 42(8):1595–1599, 1920. doi:10.1021/ja01453a010. 2.1

[LP06] Yueheng Lan and Garegin A. Papoian. The interplay between discrete noise and nonlinear chemical kinetics in a signal amplification cascade. *J. Chem. Phys.*, 125(15):154901, October 2006. doi:10.1063/1.2358342. 1.1

[LPB+06] T. Lipniacki, P. Paszek, A.R. Brasier, B. Luxon, and M. Kimmel. Transcriptional stochasticity in gene expression. *J. Theor. Biol.*, 238:348–367, 2006. doi:10.1016/j.jtbi.2005.05.032. 1.1

[MA04] Yoshihiro Morishita and Kazuyuki Aihara. Noise-reduction through interaction in gene expression and biochemical reaction processes. *J. Theor. Biol.*, 228(3):315–325, June 2004. doi:10.1016/j.jtbi.2004.01.007. 1.1

[Man07] Nikos V. Mantzaris. From single-cell genetic architecture to cell population dynamics: quantitatively decomposing the effects of different population heterogeneity sources for a genetic network with positive feedback architecture. *Biophys. J.*, 92(12):4271–288, Jun 2007. doi:10.1529/biophysj.106.100271. 1.1

[Mat07] The MathWorks. Matlab R2007b [online]. 2007. Available from: www.mathworks.com. 4.2.5

[MK06] B. Munsky and M. Khammash. A reduced model solution for the chemical master equation arising in stochastic analyses of biological networks. In *Decision and Control, 2006 45th IEEE Conference on*, pages 25–30, 2006. 2.6

[Mor07] David O. Morgan. *The Cell Cycle: Principles of Control*. Primers in Biology. New Science Press, 2007. Available from: http://www.new-science-press.com/browse/cellcycle. 4.2

[Mor08] Robert G. Mortimer. *Physical Chemistry*. Elsevier, 3rd edition, 2008. 2.2, 2.2

[NCNG+98] Béla Novák, Attila Csikasz-Nagy, Bela Gyorffy, Kathy Chen, and John J. Tyson. Mathematical model of the fission yeast cell cycle with checkpoint controls at the G1/S, G2/M and metaphase/anaphase transitions. *Biophys. Chem.*, 72(1-2):185–200, May 1998. doi:10.1016/S0301-4622(98)00133-1. 4.2

[NCT05] Béla Novák, Katherine Chen, and John Tyson. Systems biology of the yeast cell cycle engine, 2005. doi:10.1007/b137123. 1.2

[NPCT01] Béla Novák, Zsuzsa Pataki, Andrea Ciliberto, and John J. Tyson. Mathematical model of the cell division cycle of fission yeast. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(1):277–286, 2001. doi:10.1063/1.1345725. 1.2, 4, 4.2, 4.2.1, 4.2, 4.2, 4.2.1, 4.2.2, 4.2.5

[NT03] Béla Novák and John J. Tyson. Modelling the controls of the eukaryotic cell cycle. *Biochem. Soc. Trans.*, 31(Pt 6):1526–1529, 2003. Available from: http://www.biochemsoctrans.org/bst/031/bst0311526.htm. 4.2

[Nur00] Paul Nurse. A Long Twentieth Century of the Cell Cycle and Beyond. *Cell*, 100(1):71–78, January 2000. doi:10.1016/S0092-8674(00)81684-0. 4.2

[OTL+04] Ertugrul M. Ozbudak, Mukund Thattai, Han N. Lim, Boris I. Shraiman, and Alexander Van Oudenaarden. Multistability in the lactose utilization network of Escherichia coli. *Nature*, 427(6976):737–740, February 2004. doi:10.1038/nature02298. 1.1

[Pah08] Jürgen Pahle. *Stochastic simulation and analysis of biochemical networks*. PhD thesis, Humboldt-Universität zu Berlin, 2008. 1.2, 2, 2.5

[Pas07] P. Paszek. Modeling stochasticity in gene regulation: Characterization in the terms of the underlying distribution function. *Bull Math Biol*, 69:1567–1601, March 2007. doi:10.1007/s11538-006-9176-7. 1.1

[Pau04] J. Paulsson. Summing up the noise. *Nature*, 427:415–418, 2004. doi:10.1038/nature02257. 1.1

[Pau05] J. Paulsson. Models of stochastic gene expression. *Phys. Life Rev.*, 2:157–75, June 2005. doi:10.1016/j.plrev.2005.03.003. 1.1, 1.1, 1.2, 3.4, 4.2.5

[PBE00] J. Paulsson, O.G. Berg, and M. Ehrenberg. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. U. S. A.*, 97:7148–7153, 2000. doi:10.1073ypnas.110057697. 1.1, 1.1

[PE00] J. Paulsson and M. Ehrenberg. Random signal fluctuations can reduce random fluctuations in regulated components of chemical regulatory networks. *Phys. Rev. Lett.*, 84:5447–5450, 2000. doi:10.1103/PhysRevLett.84.5447. 1.1

[PE01] Johan Paulsson and Måns Ehrenberg. Noise in a minimal regulatory network: plasmid copy number control. *Quarterly Reviews Of Biophysics*, 34(1):1–59, February 2001. doi:10.1017/S0033583501003663. 1.1

[PE06] Johan Paulsson and Johan Elf. Stochastic Modeling of Intracellular Kinetics. In Zoltan Szallasi, J$^{'}$ org Stelling, and Vipul Periwal, editors, *System Modeling in Cellular Biology*, pages 149–176. The MIT Press, 2006. 1.1, 3.3

[PMK06] Slaven Peles, Brian Munsky, and Mustafa Khammash. Reduction and solution of the chemical master equation using time scale separation and finite state projection. *J. Chem. Phys.*, 125(20):204104, Nov 2006. doi:10.1063/1.2397685. 2.6

[PP01] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, fourth edition, 2001. 2

[PvO05] Juan M. Pedraza and Alexander van Oudenaarden. Noise Propagation in Gene Networks. *Science*, 307(5717):1965–1969, 2005. doi:10.1126/science.1109090. 1.1, 1.1

[Qia02] Hong Qian. From discrete protein kinetics to continuous Brownian dynamics: A new perspective. *Protein Sci.*, 11(1):1–5, 2002. Available from: http://www3.interscience.wiley.com/journal/121601869/abstract. 1.1

[Rén53] Alfréd Rényi. On the theory of order statistics. *Acta Mathematica Hungarica*, 4(3):191–231, September 1953. doi:10.1007/BF02127580. 2.3

[RO05] Jonathan M. Raser and Erin K. O'Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309(5743):2010–2013, September 2005. doi:10.1126/science.1105891. 1.1, 1.1, 1.2

[RWA02] Christopher V. Rao, Denise M. Wolf, and Adam P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature*, 420(6912):231–237, November 2002. doi:10.1038/nature01258. 1.1

[Sch72] F. Schlögl. Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik A Hadrons and Nuclei*, 253(2):147–161, April 1972. doi:10.1007/BF01379769. 2.1

[SCNG$^+$00] Akos Sveiczer, Attila Csikasz-Nagy, Bela Gyorffy, John J. Tyson, and Béla Novák. Modeling the fission yeast cell cycle: Quantized cycle times in wee1-cdc25Delta mutant cells. *Proc. Natl. Acad. Sci. U. S. A.*, 97(14):7865–7870, 2000. doi:10.1073/pnas.97.14.7865. 4.2, 4.2.1, 4.2.5, 4.2.5, 4.2.5

[SI05] M. Scott and B. P. Ingalls. Using the linear noise approximation to characterize molecular noise in reaction pathways. In *Proceedings of the AIChE Conference on Foundations of Systems Biology in Engineering (FOSBE)*, Santa Barbara, California, August 2005. arXiv:http://www.math.uwaterloo.ca/~bingalls/Pubs/ScottFOSBE.pdf. 1.2

[SIK06] Matthew Scott, Brian Ingalls, and Mads Kaern. Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(2):026107, June 2006. doi:10.1063/1.2211787. 1.2

[Sin53] K. Singer. Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 15(1):92–106, 1953. Available from: `http://www.jstor.org/pss/2983726`. 1.2

[SK04] H. El Samad and M. Khammash. Intrinsic noise rejection in gene networks by regulation of stability. In *First International Symposium on Control, Communications and Signal Processing*, pages 187–190, 2004. `doi:10.1109/ISCCSP.2004.1296252`. 1.1

[SN02] A. Sveiczer and B. Novák. Regularities and irregularities in the cell cycle of the fission yeast, Schizosaccharomyces pombe (a review), January 2002. 4.2, 4.2.1, 4.2.5, 4.2.5, 4.2.5, 4.8

[SNM96] A. Sveiczer, B. Novák, and J.M. Mitchison. The size control of fission yeast revisited. *J. Cell Sci.*, 109(12):2947–2957, 1996. Available from: `http://jcs.biologists.org/cgi/content/abstract/109/12/2947`. 4.2, 4.2.1, 4.2.5, 4.6, 4.2.5, 4.2.5, 4.8

[SOS08] Vahid Shahrezaei, Julien F. Ollivier, and Peter S. Swain. Colored extrinsic fluctuations and stochastic gene expression. *Mol Syst Biol*, 4:196, May 2008. `doi:10.1038/msb.2008.31`. 1.2

[SPA05] Michael Samoilov, Sergey Plyasunov, and Adam P. Arkin. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. U. S. A.*, 102(7):2310–2315, February 2005. `doi:10.1073/pnas.0406841102`. 1.1

[Ste04] Ralf Steuer. Effects of stochasticity in models of the cell cycle: from quantized cycle times to noise-induced oscillations. *J. Theor. Biol.*, 228(3):293–301, June 2004. `doi:10.1016/j.jtbi.2004.01.012`. 1.2, 4.2, 4.2.3, 4.2.5

[STN01] Akos Sveiczer, John J. Tyson, and Béla Novák. A stochastic, molecular model of the fission yeast cell cycle: role of the nucleocytoplasmic ratio in cycle time regulation. *Biophys. Chem.*, 92(1-2):1–15, September 2001. `doi:10.1016/S0301-4622(01)00183-1`. 4.2, 4.2.1, 4.2.5

[SU08] Tatsuo Shibata and Masahiro Ueda. Noise generation, amplification and propagation in chemotactic signaling systems of living cells. *Biosystems*, 93(1-2):126–132, 2008. `doi:10.1016/j.biosystems.2008.04.003`. 1.1

[Tan08] Moxun Tang. The mean and noise of stochastic gene transcription. *J. Theor. Biol.*, 253:271–280, 2008. `doi:10.1016/j.jtbi.2008.03.023`. 1.2

[TCN03] John J. Tyson, Katherine C. Chen, and Béla Novák. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.*, 15(2):221–231, April 2003. `doi:10.1016/S0955-0674(03)00017-6`. 4.1

[TCNN02] John J. Tyson, Attila Csikasz-Nagy, and Béla Novák. The dynamics of cell cycle regulation. *BioEssays*, 24(12):1095–1109, 2002. `doi:10.1002/bies.10191`. 1.2, 4.2

[TGOS08] Marc Turcotte, Jordi Garcia-Ojalvo, and Gürol M. Süel. A genetic timer through noise-induced stabilization of an unstable state. *Proceedings of the National Academy of Sciences*, 105(41):15732–15737, 2008. arXiv:http://www.pnas.org/content/105/41/15732.full.pdf+html, doi:10.1073/pnas.0806349105. 1.1

[TJD05] Yi Tao, Yuting Jia, and T. Gregory Dewey. Stochastic fluctuations in gene expression far from equilibrium: Omega expansion and linear noise approximation. *J. Chem. Phys.*, 122(12):124108, March 2005. doi:10.1063/1.1870874. 1.1

[TK07] Y. Togashi and K. Kaneko. Switching dynamics in reaction networks induced by molecular discreteness. *Journal Of Physics-Condensed Matter*, 19(6):065150, February 2007. doi:10.1088/0953-8984/19/6/065150. 1.1

[TSB04] T. E. Turner, S. Schnell, and K. Burrage. Stochastic approaches for modelling in vivo reactions. *Comput. Biol. Chem.*, 28(3):165–178, July 2004. doi:10.1016/j.compbiolchem.2004.05.001. 1.2, 2, 2.5

[TvO02] Mukund Thattai and Alexander van Oudenaarden. Attenuation of noise in ultrasensitive signaling cascades. *Biophys. J.*, 82(6):2943–2950, June 2002. doi:10.1016/S0006-3495(02)75635-X. 1.1

[UIYS06] Tomohiro Ushikubo, Wataru Inoue, Mitsumasa Yoda, and Masaki Sasai. Testing the transition state theory in stochastic dynamics of a genetic switch. *Chem. Phys. Lett.*, 430(1-3):139–143, October 2006. doi:10.1016/j.cplett.2006.08.114. 1.1

[UW07] M. Ullah and O. Wolkenhauer. Family tree of Markov models in systems biology. *IET Systems Biology*, 1(4):247–254, 2007. doi:10.1049/iet-syb:20070017. 2.7, 2.2

[VKBL02] José M. G. Vilar, Hao Y. Kueh, Naama Barkai, and Stanislas Leibler. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci. U. S. A.*, 99(9):5988–5992, Apr 2002. doi:10.1073/pnas.092133899. 1.1

[Vol26] Vito Volterra. Fluctuations in the abundance of a species considered mathematically. *Nature*, 118:558–560, 1926. doi:10.1038/119012b0. 2.1

[Wil06] Darren J. Wilkinson. *Stochastic Modelling for Systems Biology*, volume 11 of *Mathematical & Computational Biology*. Chapman & Hall/CRC, April 2006. 1.1, 2

[Wil09] Darren J. Wilkinson. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat Rev Genet*, 10(2):122–133, February 2009. doi:10.1038/nrg2509. 1.1

[Wri04] Margaret Robson Wright. *Introduction to Chemical Kinetics*. Wiley-Interscience, 2004. 2.2

[WUKC04] Olaf Wolkenhauer, Mukhtar Ullah, Walter Kolch, and Kwang-Hyun Cho. Modeling and simulation of intracellular dynamics: choosing an appropriate framework. *IEEE Trans Nanobioscience*, 3(3):200–207, September 2004. doi:10.1109/TNB.2004.833694. 2.4

[YJT+08] Ming Yi, Ya Jia, Jun Tang, Xuan Zhan, Lijian Yang, and Quan Liu. Theoretical study of mesoscopic stochastic mechanism and effects of finite size on cell cycle of fission yeast. *Physica A: Statistical Mechanics and its Applications*, 387(1):323–334, January 2008. doi:10.1016/j.physa.2007.07.018. 4.2, 4.2.1, 4.2.3, 4.2.5

[YUIS07] Mitsumasa Yoda, Tomohiro Ushikubo, Wataru Inoue, and Masaki Sasai. Roles of noise in single and coupled multiple genetic oscillators. *J. Chem. Phys.*, 126(11):115101, Mar 2007. doi:10.1063/1.2539037. 1.1

[ZHCN07] Judit Zamborszky, Christian I. Hong, and Attila Csikasz Nagy. Computational Analysis of Mammalian Cell Division Gated by a Circadian Clock: Quantized Cell Cycles and Cell Size Control. *J Biol Rhythms*, 22(6):542–553, 2007. doi:10.1177/0748730407307225. 1.2

[ZYDQ06] Yuping Zhang, Huan Yu, Minghua Deng, and Minping Qian. Nonequilibrium Model for Yeast Cell Cycle, 2006. doi:10.1007/11816102_84. 1.1

Publications

The work carried out in this thesis has led to the following publications:

M.Ullah, O.Wolkenhauer. Investigating the two-moment characterisation of subcellular biochemical networks. *Revised manuscript submitted to JTB on 31 Jan, 2009*

We extend previous derivations of 2MA by allowing a) non-elementary reactions and b) relative concentrations. Then we investigate the applicability of the 2MA approach to the well established fission yeast cell cycle model. Our analytical model reproduces the clustering of cycle times observed in experiments. This is explained through multiple resettings of MPF, caused by the coupling between mean and (co)variance, near the G2/M transition.

M.Ullah, O.Wolkenhauer. Family tree of Markov models in systems biology. *IET Systems Biology*, 1(4): 247-254, 2007

Motivated by applications in systems biology, a probabilistic framework based on Markov processes is proposed to represent intracellular processes. The formal relationships between different stochastic models referred to in the systems biology literature are reviewed. As part of this review, a novel derivation of the differential Chapman–Kolmogorov equation for a general multidimensional Markov process made up of both continuous and jump processes, is presented. As a result, a 'family tree' for stochastic models in systems biology is sketched, providing explicit derivations of their formal relationship and clarifying assumptions involved.

O.Wolkenhauer and M.Ullah. All Models are Wrong. *In Towards a Philosophy of Systems Biology.* F.Boogerd, F.Bruggeman, J.Hofmeyr, H.Westerhoff (eds). 163-179, 2007. ISBN 978-0-444-52085-2 (Elsevier)

We discuss Robert Rosen's critique of mathematical modelling of complex (biological) systems. The aim is to provide a concise summary of his distinction between analytical and synthetic models and to renew the interest in Rosen's work.

M.Ullah, H.Schmidt, K-H.Cho and O.Wolkenhauer. Deterministic Modelling and Stochastic Simulation of Pathways using MATLAB. *IEE Proc Systems Biology*, Vol.153, No.2, 53-60, March 2006.

We describe a collection of MATLAB functions to construct and solve ODEs for deterministic simulation and to implement realisations of CMEs for stochastic simulation using advanced MATLAB coding (Release 14). The program was successfully applied to pathway models from the literature for both cases. The results were compared to implementations

using alternative tools for dynamic modelling and simulation of biochemical networks. The aim is to provide a concise set of MATLAB functions that encourage the experimentation with systems biology models.

O.Wolkenhauer, M.Ullah, P.Wellstead, K.H.Cho. The Dynamic Systems Approach to Control and Regulation of IntraCellular Networks. *FEBS Letters*, 579 (2005), 1846-1853.

The aim of the present text is to review the systems and control perspective of dynamic systems. We review the case for differential equation models as a 'natural' representation of causal entailment in pathways. Block-diagrams, commonly used in the engineering sciences, are introduced and compared to pathway maps. Using simple examples, we show how biochemical reactions are modelled in the dynamic systems framework and visualized using block-diagrams.

O.Wolkenhauer, S.Sreenath, P.Wellstead, M.Ullah, K.H.Cho. A Systems- and Signal-Oriented Approach to IntraCellular Dynamics. *Biochemical Society Transactions*, Vol.33, Part 3, 507-515, 2005.

In this article, the framework provided by systems biology is used to argue that the same can be true for molecular biology. In particular, and using basic modular methods of mathematical modelling which are standard in control theory, a set of dynamic models is developed for some illustrative cell signalling processes. These models, supported by recent experimental evidence, are used to argue that a control theoretical approach to the mechanisms of feedback in intracellular signalling is central to furthering our understanding of molecular communication. As a specific example, a MAPK (mitogen-activated protein kinase) signalling pathway is used to show how potential feedback mechanisms in the signalling process can be investigated in a simulated environment.

O.Wolkenhauer, M.Ullah, W.Kolch, K.-H.Cho. Modelling and Simulation of IntraCellular Dynamics: Choosing an Appropriate Framework. *IEEE Transactions on NanoBioScience*, Vol. 3, No. 3, 200-207, September 2004.

We discuss in this paper the relationship between the stochastic and deterministic representations of biochemical reaction networks. Toward this end, we provide a novel compact derivation for the stochastic rate constant that forms the basis of the popular Gillespie algorithm. Comparing the mathematical basis of the two popular conceptual frameworks of generalized mass action models and the chemical master equation, we argue that some of the arguments that have been put forward are ignoring subtle differences and similarities that are important for answering the question in which conceptual framework one should investigate intracellular dynamics.

Selbständigkeitserklärung

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst habe. Es wurden ausschließlich die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht.

Rostock, 6. März 2009

Mukhtar Ullah

Curriculum Vitae

Personal Information

Forname/Surname	Mukhtar Ullah
Nationality	Pakistani
Date of birth	16 April 1975
Place of birth	Malakand, Pakistan

Education

Sep 1991 - Sep 1993	FSc. (with distinction) Pre-Engineering Islamia College, Peshawar, Pakistan
Jan 1994 - Feb 1999	BSc. (with distinction) Electrical Engineering NWFP UET Peshawar, Pakistan
Sep 2001 - Sep 2002	MSc. (with distinction) Advanced Control & Systems Engineering UMIST, Manchester, United Kingdom

Employment

Sep 2000 - Sep 2001	Teaching and Research Associate UET Taxila, Pakistan
Oct 2003 - present	Research Associate Chair in Systems Biology and Bioinformatics Department of Computer Science, University of Rostock

Theses for the doctoral research

1. The discrete and random occurrence far from thermodynamic equilibrium of chemical reactions, and low copy numbers of chemical species, in single cells necessitate stochastic approaches for correct system descriptions.
2. This work presents a stochastic framework for modelling subcellular biochemical systems. We make an effort to show how the notion of propensity, the chemical master equation and the stochastic simulation algorithm arise as consequences of the Markov property. This connection is not obvious from the relevant literature in systems biology.
3. We sketch the formal relationships between various stochastic models referred to in the systems biology literature.
4. The central theme of the present work is the two-moment approximation (2MA) as a bridge between deterministic and stochastic approaches. The 2MA combines the intuition of derministic models with the represention of noise and variability. In contrast to other stochastic approaches, an analytical 2MA model allows us to study the coupling of mean and co-variance.
5. We introduce the 2MA approach for modelling subcellular biochemical systems and develop extensions to allow a) non-elementary reactions and b) relative concentrations.
6. We investigate the applicability of the 2MA approach with the Tyson-Novák model for the fission yeast cell cycle model.
7. Our analytical model reproduces the clustering of cycle times observed in experiments. This is explained through multiple resettings of a protein called MPF, caused by the coupling between mean and (co)variance, near the G2/M transition.
8. One notable aspect of our analytical model is that, although it describes the average of an ensemble, it reproduces enough variability among cycles to reproduce the clustering of cycle times observed in experiments on double mutants.
9. The 2MA approach can thus infer new properties in a single simulation run that is neither possible with deterministic approach (no representation of noise) nor stochastic approach (requiring many simulation runs).