ON MOMENTS AND TIMING – STOCHASTIC ANALYSIS OF BIOCHEMICAL SYSTEMS

by

Khem Raj Ghusinga

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Summer 2018

ON MOMENTS AND TIMING – STOCHASTIC ANALYSIS OF BIOCHEMICAL SYSTEMS

by

Khem Raj Ghusinga

Approved: _____

Kenneth E. Barner, Ph.D.
Chair of the Department of Electrical and Computer Engineering

Approved: _____

Babatunde Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____

Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Abhyudai Singh, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Ryan Zurakowski, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Gonzalo R. Arce, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Pak–Wing Fok, Ph.D.
Member of dissertation committee

ACKNOWLEDGEMENTS

I express my deepest gratitude to my advisor, Prof. Abhyudai Singh, as this thesis would not have been possible without his constant guidance and support. Thanks are also due to Prof. Ryan Zurakowski, Prof. Gonzalo Arce, and Prof. Pak-Wing Fok for being part of my PhD committee and for numerous discussions on this research and beyond.

I am grateful to Prof. Andrew Lamperski, Prof. John J. Dennehy, Prof. Pak–Wing Fok, and Prof. James Umen with whom I have collaborated on research projects. Working with them has exposed me to different approaches to research and has helped me formulate my own approach.

I acknowledge all members of the Singh lab, both present and past, for their regular feedback and support. My colleagueship with César and Mohammad developed into close friendships that I would cherish for lifetime.

I thank all my friends who supported me in numerous ways throughout. My roommates provided a perfect environment to take my mind off research. Nistha merits a special mention for being my strongest critic as well as my ultimate supporter. Finally, I am grateful to my family for their unwavering support. I would strive to be at least a fraction as caring as they have been to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

At the level of individual living cells, key species such as genes, mRNAs, and proteins are typically present in small numbers. Consequently, the biochemical reactions involving these species are inherently noisy and result in considerable cell-to-cell variability. This thesis outlines two mathematical tools to quantify stochasticity in these biochemical reaction systems: (i) a novel computational method that provides exact lower and upper bounds on statistical moments of population counts of important species, and (ii) a first-passage time framework to study noise in the timing of a cellular event that occurs when population count of an underlying regulatory protein attains a critical threshold.

The method to compute bounds on moments builds upon the well-known linear dynamical system that describes the time evolution of statistical moments. However, except for some ideal cases, this dynamical system is not closed in the sense that lower-order moments depend upon some higher-order moments. To overcome this issue, our method exploits the fact that statistical moments of a random variable must satisfy constraints that are compactly represented through the positive semidefiniteness of moment matrices. We find lower and upper bounds on a moment of interest via a semidefinite program that includes linear constraints obtained from moment dynamics, along with semidefinite constraints on moment matrices. We further show that these bounds improve as the size of the semidefinite program is increased by including dynamics of more moments as well as constraints involving them. We also extend the scope of this method for stochastic hybrid systems, which are a more general class of stochastic systems that integrate discrete and continuous dynamics.

The second tool proposed in this thesis–a first-passage time framework to study event timing–is based on the premise that several cellular events in living cells occur

upon attainment of critical levels by corresponding regulatory proteins. Two particular examples that we study here are the lysis of a bacterial cell infected by the virus bacteriophage λ and the cell-division in exponentially growing bacterial cells. We provide analytical calculations for the first-passage time distribution and its moments for both these examples. We show that the first-passage time statistics can be used to explain several experimentally observed behaviors in both these systems. Finally, the thesis discusses potential directions for future research.

Chapter 1

INTRODUCTION

Stochasticity or noise is an integral aspect of biochemical systems. For example, genetically identical cells are known to exhibit considerable variability in their behavior. This variability can in principle stem from fluctuations in external environment of the cells (temperature, nutrients, stresses, etc.) as well as in intrinsic factors. However, the cell-to-cell variability is present even in a controlled and supposedly homogenous environment. With recent progress in single-cell and single-molecule technologies, it has become apparent that the inherent probabilistic nature of biochemical reactions, coupled with the fact that important species such as genes, mRNAs, etc. occur in individual cells at low copy numbers, leads to noise in gene-product levels across isogenic cells [5–16]. This noise then manifests as variability in biological functions [17–32].

Prevalence of noise in biochemical systems raises several fundamental questions. For example, how cells behave reliably despite the presence of noise at the molecular and cellular levels? Is it possible to exploit noise for certain cellular functions? Have single-cell and more complex organisms evolved to exploit noise to enhance performance? These questions have been widely investigated in the last decade and it has been shown that cells use a variety of regulatory mechanisms to suppress or exploit fluctuations in gene-product levels [33–43]. Interestingly, the noise in gene-product levels can also manifest as noise in other important variables, such as timing of cellular events that occur up on accumulation of a regulatory proteins up to critical threshold levels. The questions of reliable performance in presence of noise or its utilization to a cell's advantage can also be asked in context of timing of events [44, 45].

Given how important the noise in gene-product levels (and consequently in other variables such as timing of cellular events) is, quantification of noise is extremely important for understanding the cellular function. Towards that end, this thesis develops two mathematical tools:

- A novel computational method that provides exact lower and upper bounds on statistical moments of gene-product levels. These tools are later extended to stochastic hybrid systems which are a much broader class of stochastic systems.
- A first-passage time framework to study noise in timing of cellular events that occur once levels of associated regulatory proteins achieve some critical thresholds.

The first tool proposed above is described in Chapter 2. It utilizes the well-known linear dynamical system that governs time evolution of moments. It turns out that when the biochemical system contains nonlinearities, the linear dynamical system is an infinite-hierarchical system of coupled differential equations. Computing, or even estimating, moments of species level in such systems has implications in not only quantitative understanding of these systems, but also in parameter inference [46–50]. Furthermore, accurate estimation of moments are desirable for identification of molecular underpinnings of biochemical processes [9]. Whereas current methods to approximate moments are based on ad-hoc assumptions [48, 51–73], we propose a method to obtain both lower and upper bounds on moments of a biochemical system without any assumption. The method uses the steady-state moment equations along with semidefinite constraints that required to be satisfied by moments of a random variable. These inequalities are constructed from positive semidefinite constraints on moments of a positive random variable. We show that not only one can obtain upper and lower bounds on a given stationary moment, but also both upper and lower bounds improve considerably as one uses more moment equations. Later in Chapter 3, we extend the scope of the method to stochastic hybrid systems (SHSs) that integrate both continuous dynamics with discrete events. We show that our method extends to both polynomial SHSs and rational SHSs. Furthermore, we also discuss an alternate method to estimate moments for simple SHSs that exploits the positivity of the characteristic function.

With regards to the second tool proposed here, we formulate event timing as the first-passage time for a protein's level to cross a threshold. Although we study two particular problems here – the lysis time of an *E. coli* cell infected by a bacteriophage λ (Chapter 4) and cell division time in bacteria (Chapter 5) – there are several other examples in context of development [74–77], cell-cycle control [78–82], cell differentiation [83, 84], sporulation [85, 86], apoptosis [87–89], etc. wherein the proposed framework could be useful. We particularly ask how different parameters of interest shape the statistics of event timing. Moreover, in context of bacteriophage λ, lysis time is presumably scheduled so as to minimize noise in timing, and therefore also examine how cells would regulate gene expression to minimize noise in timing around a given mean time. For both examples, we discuss how the proposed framework explains important experimentally observed characteristics of event timing.

1.1 Notation

For stochastic processes and their moments, we omit explicit dependence on time unless it is not clear from the context. Inequalities for vectors are element-wise. Random variables are denoted in bold. The n-dimensional Euclidian space is denoted by $\mathbb{R}^n$. The set of non-negative integers is denoted by $\mathbb{N}$ whereas set of complex numbers is represented by $\mathbb{C}$. $\mathbb{E}(\boldsymbol{x})$ or $\langle \boldsymbol{x} \rangle$ is used for expectation of a random variable $\boldsymbol{x}$.

1.2 Publication List

This thesis is based on the material in the following publications:

[90] (Chapter 2) K. Ghusinga, C. Vargas-Garcia, A. Lamperski, and A. Singh. Exact lower and upper bounds on stationary moments in stochastic biochemical systems. Physical Biology, 14:04LT01, 2017

[91] (Chapter 3 & 4) A. Lamperski, K. Ghusinga, and A. Singh. Analysis and Control of Stochastic Systems using Semidefinite Programming over Moments. Under Review, 2017

[92] (Chapter 3) K. Ghusinga, A. Lamperski, and A. Singh. Moment Analysis of Stochastic Hybrid Systems using Semidefinite Programming. Under Review, 2018

[93] (Chapter 3) A. Lamperski, K. Ghusinga, and A. Singh. Stochastic optimal control using semidefinite programming for moment dynamics. In 55th IEEE Conference on Decision and Control, pages 1990–1995, 2016

[94] (Chapter 3) K. Ghusinga, A. Lamperski, and A. Singh. Estimating Stationary Characteristic Functions of Stochastic Systems via Semidefinite Programming. In European Control Conference, to appear, 2018

[95] (Chapter 4) K. Ghusinga, J. Dennehy, and A. Singh. First-passage time approach to controlling noise in the timing of intracellular events. Proceedings of the National Academy of Sciences, 114:693–698, 2017

[96] (Chapter 4) K. Ghusinga and A. Singh. Effect of gene-expression bursts on stochastic timing of cellular events. In 2017 American Control Conference (ACC), pages 2118–2123, 2017

[97] (Chapter 4) K. Ghusinga and A. Singh. Optimal regulation of protein degradation to schedule cellular events with precision. In 2016 American Control Conference (ACC), pages 424–429, 2016

[98] (Chapter 4) K. Ghusinga, P.-W. Fok, and A. Singh. Optimal auto-regulation to minimize first-passage time variability in protein level. In 2015 American Control Conference (ACC), pages 4411–4416, 2015

[99] (Chapter 4) K. Ghusinga and A. Singh. Theoretical predictions on the first-passage time for a gene expression model. In 54th IEEE Conference on Decision and Control, pages 3864–3869, 2015

[100] (Chapter 4) K. Ghusinga and A. Singh. First-passage time calculations for a gene expression model. In 53rd IEEE Conference on Decision and Control, pages 3047–3052, 2014

[101] (Chapter 5) K. Ghusinga, C. Vargas-Garcia, and A. Singh. A mechanistic stochastic framework for regulating bacterial cell division. Scientific Reports, 6:30229, 2016

Chapter 2

BOUNDING MOMENTS OF BIOCHEMICAL SYSTEMS

In a well-mixed biochemical reaction system, the reactions are inherently stochastic owing to the perpetual random motion at the molecular level. These stochastic effects are particularly relevant when the species are present at low molecular counts. Mathematical characterization of such systems is done by employing the chemical master equation (CME) [102–104]. To formally write the CME, consider a system of n species $X_j, j = \{1, 2, \ldots, n\}$ and denote the state of the system by a vector $\boldsymbol{x}(t) = \begin{bmatrix} \boldsymbol{x}_1(t) & \boldsymbol{x}_2(t) & \ldots & \boldsymbol{x}_n(t) \end{bmatrix}^\top$, where $\boldsymbol{x}_j(t)$ represents the population of X_j at time t. The time evolution of the probability, $\mathbb{P}\left(\boldsymbol{x}(t)\right)$, that the system is in state $\boldsymbol{x}(t)$ follows the following differential equation

$$\frac{d\mathbb{P}\left(\boldsymbol{x}(t)\right)}{dt} = \sum_{i=1}^{S} f_i(\boldsymbol{x}(t) - \eta_i)\mathbb{P}\left(\boldsymbol{x}(t) - \eta_i\right) - \sum_{i=1}^{S} f_i(\boldsymbol{x}(t))\mathbb{P}\left(\boldsymbol{x}(t)\right). \tag{2.1}$$

Here S represents the number of reactions $\mathcal{R}_i$, $i \in \{1, 2, \ldots, S\}$ through which the species interact. The term $f_i(\boldsymbol{x}(t))$ denotes the propensity of the i^{th} reaction $\mathcal{R}_i$ in the sense that the probability that it occurs in an infinitesimal small time-interval $[t, t+dt)$ is given by $f_i(\boldsymbol{x}(t))dt$. Upon occurrence of $\mathcal{R}_i$, the state of the system is transitioned to $\boldsymbol{x} + \eta_i$ where η_i is the stoichiometry vector that describes the change in population.

It turns out that the CME is analytically intractable except for few special cases, and generally requires considerable computation effort if solved numerically [105–114]. The computational cost tends to become prohibitive if one is interested in studying the long time (i.e., stationary or steady-state) behavior of the system. Perhaps a reasonable goal is to determine a few lower order moments (such as mean, variance, etc.) of different species. Not only is moment computation of primary importance for

many purposes, but it can also be used to infer useful information about the probability density function using the tools such as the Chebyshev's inequality [115], moment-based reconstruction of the probability density function [116], etc.

The time evolution of moments of a biochemical system is governed via a system of differential equations which can be obtained from the CME [117–119]. Given a vector $m = \begin{bmatrix} m_1 & m_2 & \dots m_n \end{bmatrix}^\top$ of n non-negative integers, a statistical moment of $\boldsymbol{x}$ is defined as $\langle \boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \cdots \boldsymbol{x}_n^{m_n} \rangle$ where the sum $\sum_{j=1}^{n} m_j$ is referred to as the order of the moment. Using a short-hand $\boldsymbol{x}^{[m]} := \boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \cdots \boldsymbol{x}_n^{m_n}$, the time derivative of $\left\langle \boldsymbol{x}^{[m]} \right\rangle$ obtained from the CME is given by [117–119]

$$\frac{d\left\langle \boldsymbol{x}^{[m]} \right\rangle}{dt} = \left\langle \sum_{i=1}^{S} f_i(\boldsymbol{x}) \left((\boldsymbol{x} + \eta_i)^{[m]} - \boldsymbol{x}^{[m]} \right) \right\rangle. \tag{2.2}$$

Typically the propensity functions $f_i(\boldsymbol{x})$ for mass action kinetics are taken as finite polynomials in $\boldsymbol{x}$ [102–104]. In that case, it follows from (2.2) that if all moments up to order M are stacked in a vector $\mathcal{X}$ (including the zeroth order moment which is equal to 1), then the time evolution of $\mathcal{X}$ is given by

$$\frac{d\mathcal{X}}{dt} = A\mathcal{X} + \bar{A}\,\overline{\mathcal{X}} \tag{2.3}$$

Here $\overline{\mathcal{X}}$ is a finite vector consisting of moments of order higher than M [118]. The elements of matrices A and B depend upon reaction parameters.

When the reaction propensities are constants (zero-order reactions), linear (first-order reactions), or the system has some special structure [120], then $\bar{A} = 0$ and the moments in $\mathcal{X}$ can be determined exactly by solving (2.3). However, in general the matrix $\bar{A} \neq 0$, which implies that (2.3) represents an unclosed system of differential equations. One widely used approach for handling such cases is to employ an appropriate moment closure technique. Based on different assumptions, these techniques approximate the vector $\overline{\mathcal{X}}$ as a, possibly nonlinear, function of $\mathcal{X}$ [48,51–73]. Although presumed to be reasonably accurate, the moment closure schemes typically do not provide any mathematical guarantees on the accuracy of the approximation.

In this chapter we present an alternate method that provides both upper and lower bounds on moments. Here, instead of finding an approximation of $\overline{\mathcal{X}}$, we use the fact that at all times the elements of $\mathcal{X}$ and $\overline{\mathcal{X}}$ are moments of some probability distribution. Therefore, these elements cannot take arbitrary values and must satisfy some constraints (e.g., variance is non-negative). These constraints are compactly represented in terms of positive semidefiniteness of moment matrices [121]. We show that solutions to the moment dynamics that preserve the positive semidefiniteness of moment matrices results in the lower and upper bounds on elements of $\mathcal{X}$ at a given time. Furthermore, increasing the order of truncation, i.e., the size of $\mathcal{X}$, and correspondingly increasing the number of moment constraints often results in improvement in these bounds.

2.1 Constraints on Moments

In the proposed method, our aim is to exploit the constraints satisfied by moments of any random variable [121]. For simplicity, let us first consider the case of a scalar random variable $\boldsymbol{x}$. Suppose we construct a vector $\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \boldsymbol{x} & \boldsymbol{x}^2 & \dots \boldsymbol{x}^d \end{bmatrix}^\top$ that consists of monomials up to degree d of $\boldsymbol{x}$. Then the outer product $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top$ is positive semidefinite (denoted by $\succeq 0$), and the semidefiniteness is preserved if expectation is taken (see Appendix A). That is, we have

$$H_e = \left\langle \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle = \begin{bmatrix} 1 & \langle \boldsymbol{x} \rangle & \dots & \langle \boldsymbol{x}^d \rangle \\ \langle \boldsymbol{x} \rangle & \langle \boldsymbol{x}^2 \rangle & \dots & \langle \boldsymbol{x}^{d+1} \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle \boldsymbol{x}^d \rangle & \langle \boldsymbol{x}^{d+1} \rangle & \dots & \langle \boldsymbol{x}^{2d} \rangle \end{bmatrix} \succeq 0, \tag{2.4}$$

for all $d = \{1, 2, \dots\}$. Furthermore, if the random variable $\boldsymbol{x}$ is non-negative, which is the case when $\boldsymbol{x}$ represents the level of some biochemical species, another semidefinite constraint can be obtained as

$$H_{o_i} = \left\langle h_i(\boldsymbol{x}) \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle, \tag{2.5}$$

where $h_i(\boldsymbol{x})$ belongs to a family of non-negative polynomial functions of $\boldsymbol{x}$ that are indexed by i. For example, when $h_i(\boldsymbol{x}) = \boldsymbol{x}$, we have

$$\begin{bmatrix} \langle \boldsymbol{x} \rangle & \langle \boldsymbol{x}^2 \rangle & \dots & \langle \boldsymbol{x}^{d+1} \rangle \\ \langle \boldsymbol{x}^2 \rangle & \langle \boldsymbol{x}^3 \rangle & \dots & \langle \boldsymbol{x}^{d+2} \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle \boldsymbol{x}^{d+1} \rangle & \langle \boldsymbol{x}^{d+2} \rangle & \dots & \langle \boldsymbol{x}^{2d+1} \rangle \end{bmatrix} \succeq 0, \tag{2.6}$$

for all $d = \{1, 2, \dots\}$. This fact can be further exploited if $\boldsymbol{x}$ has bounded support. For example, if $\boldsymbol{x} \in [L, U] \subset \mathbb{R}$ then $\left\langle (\boldsymbol{x} - L)\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle \succeq 0$ and $\left\langle (U - \boldsymbol{x})\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle \succeq 0$.

These constraints can be generalized to multivariate random variables. For example, for an n-dimensional random variable $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^\top$, a matrix H_e that is analogous to the one in (2.4) can be constructed by taking expectation of outer product $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top$ where the vector $\boldsymbol{\Gamma}$ (with slight abuse of the notation) consists of all monomials of $\boldsymbol{x}$ up to order d

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \boldsymbol{x}_1 \dots \boldsymbol{x}_n & \boldsymbol{x}_1^2 & \boldsymbol{x}_1\boldsymbol{x}_2 \dots \boldsymbol{x}_1\boldsymbol{x}_n \dots \boldsymbol{x}_n^2 \dots \boldsymbol{x}_n^d \end{bmatrix}^\top. \tag{2.7}$$

Furthermore, a family of matrices analogous to H_{o_i} in (2.6) can be generated as $\left\langle h_i(\boldsymbol{x})\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle \succeq 0$. For example, when $h_i(\boldsymbol{x}) = \boldsymbol{x}_i$ then n matrices can be obtained as $\left\langle \boldsymbol{x}_i\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle \succeq 0$ for $i = 1, 2, \dots, n$.

The constraints described by (2.4)–(2.6) and their multivariate analogues essentially ensure that the higher order moments appearing in (2.3) do not take arbitrary values. To see this, we can use the fact that a matrix is positive semidefinite if and only if all of its leading principal minors are non-negative (this is known as the Sylvester's criterion [122]). Therefore, for $d = 1$, the non-negative determinants of the matrices in (2.4) and (2.6) result in

$$\left\langle \boldsymbol{x}^2 \right\rangle \geq \left\langle \boldsymbol{x} \right\rangle^2, \quad \left\langle \boldsymbol{x}^3 \right\rangle \geq \frac{\left\langle \boldsymbol{x}^2 \right\rangle^2}{\left\langle \boldsymbol{x} \right\rangle}, \tag{2.8}$$

respectively. Note that the first inequality above is nothing but the well-known inequality representing non-negativity of variance. Similarly, for $d = 2$, the determinant of the matrix in (2.4) yields

$$\left\langle \boldsymbol{x}^4 \right\rangle \geq \frac{\left\langle \boldsymbol{x}^3 \right\rangle^2 + \left\langle \boldsymbol{x}^2 \right\rangle^3 - 2\left\langle \boldsymbol{x}^3 \right\rangle \left\langle \boldsymbol{x}^2 \right\rangle \left\langle \boldsymbol{x} \right\rangle}{\left\langle \boldsymbol{x}^2 \right\rangle - \left\langle \boldsymbol{x} \right\rangle^2}. \tag{2.9}$$

In essence, these determinants for varying d allow higher-order moments to be bounded from below by nonlinear functions of the lower-order moments.

Another point to note is that the matrix $\left\langle \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle$ generates inequalities for even order moments whereas $\left\langle \boldsymbol{x}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle$ generates inequalities for odd order moments. Likewise, for the multivariate random variable $\boldsymbol{x}$, non-negativity of the principal minors of matrix $\left\langle \boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle$ gives bounds on the moments $\left\langle \boldsymbol{x}_i^{2d} \right\rangle$, $i = \{1, 2, \ldots, n\}$. As an example, for $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix}^\top$, the following is obtained for $d = 1$

$$\left\langle \begin{bmatrix} 1 & \boldsymbol{x}_1 & \boldsymbol{x}_2 \\ \boldsymbol{x}_1 & \boldsymbol{x}_1^2 & \boldsymbol{x}_1\boldsymbol{x}_2 \\ \boldsymbol{x}_2 & \boldsymbol{x}_1\boldsymbol{x}_2 & \boldsymbol{x}_2^2 \end{bmatrix} \right\rangle \succeq 0, \tag{2.10}$$

which results in $\left\langle \boldsymbol{x}_1^2 \right\rangle \geq \left\langle \boldsymbol{x}_1 \right\rangle^2$, and the following inequality bounding $\left\langle \boldsymbol{x}_2^2 \right\rangle$

$$\left\langle \boldsymbol{x}_2^2 \right\rangle \geq \frac{\left\langle \boldsymbol{x}_1\boldsymbol{x}_2 \right\rangle^2 + \left\langle \boldsymbol{x}_1^2 \right\rangle \left\langle \boldsymbol{x}_2 \right\rangle - 2\left\langle \boldsymbol{x}_1 \right\rangle \left\langle \boldsymbol{x}_2 \right\rangle \left\langle \boldsymbol{x}_1\boldsymbol{x}_2 \right\rangle}{\left\langle \boldsymbol{x}_1^2 \right\rangle - \left\langle \boldsymbol{x}_1 \right\rangle^2}. \tag{2.11}$$

The moments whose form is different from $\left\langle \boldsymbol{x}_i^{2d} \right\rangle$ can be bounded by taking expectation of $\left\langle h_i(\boldsymbol{x})\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle$ with appropriately chosen functions, h_i's. For instance, when $\boldsymbol{x}$ takes positive values, $h_i(\boldsymbol{x}) = \boldsymbol{x}_1$ gives the constraint

$$\left\langle \boldsymbol{x}_1 \begin{bmatrix} 1 & \boldsymbol{x}_1 & \boldsymbol{x}_2 \\ \boldsymbol{x}_1 & \boldsymbol{x}_1^2 & \boldsymbol{x}_1\boldsymbol{x}_2 \\ \boldsymbol{x}_2 & \boldsymbol{x}_1\boldsymbol{x}_2 & \boldsymbol{x}_2^2 \end{bmatrix} \right\rangle \succeq 0, \tag{2.12}$$

which can be used to find a lower bound on $\left\langle \boldsymbol{x}_1\boldsymbol{x}_2^2 \right\rangle$.

It should also be noted that the univariate inequalities obtained from (2.4)–(2.6) are valid for both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. Furthermore, as $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are positive random variables, additional inequalities as follows can also be written

$$\left\langle \boldsymbol{x}_1 \begin{bmatrix} 1 & \boldsymbol{x}_2 \\ \boldsymbol{x}_2 & \boldsymbol{x}_2^2 \end{bmatrix} \right\rangle \succeq 0, \quad \left\langle \boldsymbol{x}_2 \begin{bmatrix} 1 & \boldsymbol{x}_1 \\ \boldsymbol{x}_1 & \boldsymbol{x}_1^2 \end{bmatrix} \right\rangle \succeq 0. \tag{2.13}$$

These inequalities translate to lower bounds on moments $\langle \boldsymbol{x}_1^2 \boldsymbol{x}_2 \rangle$ and $\langle \boldsymbol{x}_1 \boldsymbol{x}_2^2 \rangle$. To sum up, one can write a multitude of positive semidefinite matrix constraints satisfied by the moments.

For simplicity we first illustrate the methodology to obtain bounds on stationary moments. Here we implicitly assume that at least one stationary distribution with valid moments exist. This is typically a valid assumption for biochemical systems, and the reader may refer to [123, 124] for technical details. We find that the approach is particularly well suited for stationary moments, although it is applicable to bound moments at any given time.

2.2 Bounds on Steady-State Moments

In this section, we provide a general methodology to obtain bounds on stationary moments of a biochemical system, and illustrate it using examples.

2.2.1 Method

Broadly speaking, the method utilizes the linear equations obtained from the moment dynamics in (2.3) and constraints on the higher order moments discussed in the previous section. The key steps are enumerated below

1. Obtain the system of linear equations $A\mathcal{X} + \bar{A}\overline{\mathcal{X}} = 0$ for the stationary moments.
2. Solve $A\mathcal{X} + \bar{A}\overline{\mathcal{X}} = 0$ for the elements of $\mathcal{X}$ in terms of the elements of $\overline{\mathcal{X}}$.
3. Use inequalities which bound the elements of $\overline{\mathcal{X}}$ in terms of elements of $\mathcal{X}$ in conjuction with equations in step 2. This yields inequalities for the elements of $\overline{\mathcal{X}}$.
4. Usually, there are multiple solutions; the spurious solutions can be discarded based on moment inequalities satisfied by the lower order moments.

5. Reverting to step 2, the bound on elements of $\mathcal{X}$ are readily obtained by substituting the corresponding bounds on $\overline{\mathcal{X}}$.

These steps can be easily implemented in a computational tool such as Mathematica or Matlab, and can even give analytical bounds if the problem at hand is small enough. However, it turns out that as the size of the problem grows, solving the inequalities becomes difficult for these off-the-shelf tools.

A more specialized approach is to recognize the fact that we are essentially looking for minimum or maximum value taken by a moment of interest, subject to linear constraints arising from the moment equations $A\mathcal{X} + \bar{A}\,\overline{\mathcal{X}} = 0$ and the semidefinite constraints in (2.4)–(2.6) (or their multivariate analogues). Thus, the problem could be posed as a semidefinite program [125] whereby maximizing (minimizing) a moment gives upper (lower) bound. For example, a lower bound on a moment of interest $\mu \in \mathcal{X}$ can be computed via the semidefinite program

$$\begin{aligned}
\underset{\mathcal{X},\overline{\mathcal{X}}}{\text{minimize}} \quad & \mu && \text{(2.14a)}\\
\text{subject to} \quad & 0 = A\mathcal{X} + \bar{A}\overline{\mathcal{X}} && \text{(2.14b)}\\
& H_e \succeq 0 && \text{(2.14c)}\\
& H_{o_i} \succeq 0. && \text{(2.14d)}
\end{aligned}$$

The upper bound can be computed by multiplying the objective function by -1. These semidefinite programs could be solved using specialized algorithms [125]. In what follows, we will illustrate the proposed method via examples.

2.2.2 Stochastic Logistic Growth with Constant Immigration Rate

Consider the following biochemical system, where a species X arrives in the system via two modes, a constant immigration rate k and a species dependent rate r, and each species degrades or leaves the system with a rate $\frac{r\boldsymbol{x}}{\Omega}$

$$\emptyset \xrightarrow{k} X, \quad X \xrightarrow{r} X+1, \quad X \xrightarrow{r\boldsymbol{x}/\Omega} X-1, \tag{2.15}$$

where x denotes the population level of the species. With $k = 0$, this model essentially represents a logistic growth model which is widely used to model growth of populations in ecology, and virus dynamics [59,126–128]. In the deterministic sense, the population grows with a rate r and saturates once it reaches a finite carrying capacity Ω due to resource limitations. The term k here represents a constant rate of immigration so as to avoid the extinction of the population.

In this example, there are three reactions with propensity functions and stoichiometric vectors tabulated in Table 2.1. Using (2.2), we can write the moment dynamics for a m^{th} order moment as

$$\frac{d\left\langle \boldsymbol{x}^m \right\rangle}{dt} = \left\langle (k + r\boldsymbol{x})\left((\boldsymbol{x}+1)^m - \boldsymbol{x}^m\right) + \frac{r}{\Omega}\boldsymbol{x}^2\left((\boldsymbol{x}-1)^m - \boldsymbol{x}^m\right)\right\rangle. \tag{2.16}$$

Note that these equations are not closed because of the quadratic propensity function, and time evolution of a m^{th} order moment depends upon a moment of $(m+1)^{th}$ order. Thus, if we construct a vector $\mathcal{X}$ consisting of first M moments, then $\overline{\mathcal{X}} = [\langle \boldsymbol{x}^{M+1} \rangle]$.

Table 2.1: Description of reactions for the logistic growth model

Reaction, $\mathcal{R}_i$	Stoichiometric vector, η_i	Propensity function, $f_i(\boldsymbol{x})$
$\emptyset \xrightarrow{k} X + 1$	$\begin{bmatrix}1\end{bmatrix}^\top$	k
$X \xrightarrow{r} X + 1$	$\begin{bmatrix}1\end{bmatrix}^\top$	$r\boldsymbol{x}$
$X \xrightarrow{r\boldsymbol{x}/\Omega} X - 1$	$\begin{bmatrix}-1\end{bmatrix}^\top$	$(r/\Omega)\boldsymbol{x}^2$

To see how the moment inequalities lead to bounds, we start with the simplest case of $M = 1$. From (2.45), the steady-state moment satisfies

$$\frac{d\langle \boldsymbol{x} \rangle}{dt} = k + r\langle \boldsymbol{x} \rangle - \frac{r}{\Omega}\langle \boldsymbol{x}^2 \rangle = 0. \tag{2.17}$$

Solving the above equation gives

$$\langle \boldsymbol{x} \rangle = \frac{r\langle \boldsymbol{x}^2 \rangle - k\Omega}{r\Omega}. \tag{2.18}$$

Using this with the non-negative variance inequality $\langle \boldsymbol{x}^2 \rangle \geq \langle \boldsymbol{x} \rangle^2$ gives a quadratic inequality in $\langle \boldsymbol{x}^2 \rangle$

$$r^2 \left\langle \boldsymbol{x}^2 \right\rangle^2 - \left(2kr\Omega + r^2\Omega^2\right) \left\langle \boldsymbol{x}^2 \right\rangle + k^2\Omega^2 \leq 0, \tag{2.19}$$

whose solution can be used in (2.18) to obtain bounds on $\langle x \rangle$

$$\frac{\Omega}{2} - \frac{1}{2}\sqrt{\frac{4k\Omega + \Omega^2 r}{r}} \leq \langle \boldsymbol{x} \rangle \leq \frac{1}{2}\sqrt{\frac{4k\Omega + \Omega^2 r}{r}} + \frac{\Omega}{2}. \tag{2.20}$$

The lower bound on $\langle x \rangle$ can be discarded from the fact that $\langle \boldsymbol{x} \rangle \geq 0$. Thus, we have the following lower and upper bounds on $\langle x \rangle$

$$0 \leq \langle \boldsymbol{x} \rangle \leq \frac{1}{2}\sqrt{\frac{4k\Omega + \Omega^2 r}{r}} + \frac{\Omega}{2}. \tag{2.21}$$

Next, consider the case of $M = 2$. In this case, the steady-state moment equations are given by

$$\frac{d\langle \boldsymbol{x} \rangle}{dt} = k + r \langle \boldsymbol{x} \rangle - \frac{r}{\Omega} \left\langle \boldsymbol{x}^2 \right\rangle = 0, \tag{2.22}$$

$$\frac{d\langle \boldsymbol{x}^2 \rangle}{dt} = k + \left(2k + \frac{r}{\Omega}\right) \langle \boldsymbol{x} \rangle + \left(2r + \frac{r}{\Omega}\right) \left\langle \boldsymbol{x}^2 \right\rangle - \frac{2r}{\Omega} \left\langle \boldsymbol{x}^3 \right\rangle = 0. \tag{2.23}$$

Solving these equations results in

$$\langle \boldsymbol{x} \rangle = -\frac{k\Omega^2 + k\Omega - r \left\langle \boldsymbol{x}^3 \right\rangle}{\Omega(k + \Omega r + r)}, \quad \left\langle \boldsymbol{x}^2 \right\rangle = \frac{k^2\Omega + r^2 \left\langle \boldsymbol{x}^3 \right\rangle}{r(k + \Omega r + r)}. \tag{2.24}$$

Using the above expressions in the second inequality from (2.8) leads to a quadratic inequality in $\langle \boldsymbol{x}^3 \rangle$. Substituting the solution back to expressions of $\langle \boldsymbol{x} \rangle$ and $\langle \boldsymbol{x}^2 \rangle$ results in bounds on them. Upon rejecting the spurious solutions by using inequalities $\langle \boldsymbol{x}^2 \rangle \geq \langle \boldsymbol{x} \rangle$ and $\langle \boldsymbol{x} \rangle \geq 0$, a more useful lower bound for $\langle \boldsymbol{x} \rangle$ is found as compared to the $M = 1$ case

$$\frac{1}{2}\sqrt{\frac{4k^3\Omega + k^2\Omega^2 r + 2k^2\Omega r + k^2 r}{r(k + r)^2}} + \frac{k(\Omega - 1)}{2(k + r)} \leq \langle \boldsymbol{x} \rangle \leq \frac{1}{2}\sqrt{\frac{4k\Omega + \Omega^2 r}{r}} + \frac{\Omega}{2}. \tag{2.25}$$

As the order of truncation M is increased, the same approach can be used: take steady-state equations of first M moments, use the inequality bounding the

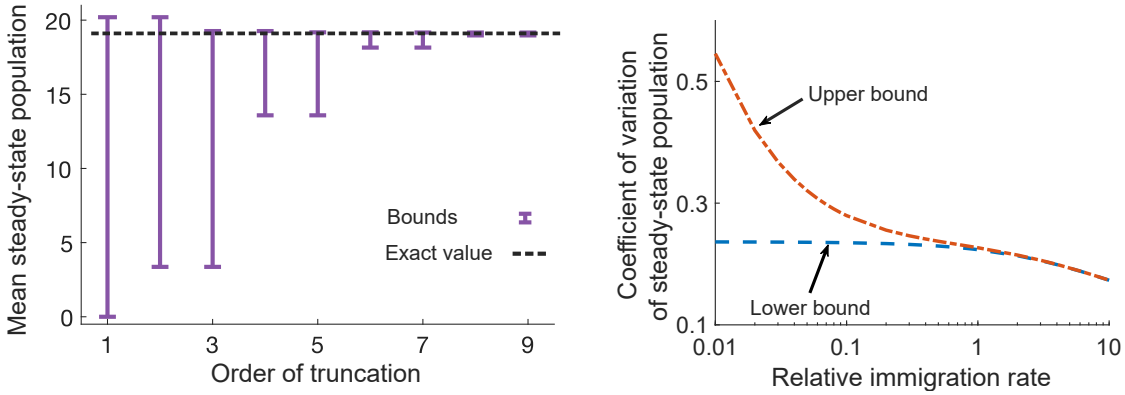

Figure 2.1: Estimated bounds on steady-state moments for the logistic growth model. *Left:* Upper and lower bounds on the second moment of the population level are shown for different orders of truncation M. As M is increased, the bounds obtained get tighter. The exact mean value of 19.149 is obtained by averaging $100,000$ MC simulations performed using SSA [1]. Parameters in (2.15) taken as $k = 1$, $r = 5$, and $\Omega = 20$. *Right:* The bounds on the coefficient of variation (standard deviation/mean) of the steady-state population level are shown as the relative immigration rate k/r is changed. Both lower and upper bounds decrease as the relative immigration rate increases. These bounds are obtained via a 5^{th} order truncation.

$(M+1)^{th}$ moment, and apply the inequalities for lower order moments to prune solutions. Though the resulting expressions do not lead to closed-form analytical bounds, numerical solutions are still possible. Interestingly, the solutions for odd values of M improve the upper bounds on the average population level whereas the solutions for even values of M improve the lower bounds (Fig. 2.4 (Left)). At $M = 9$, the lower and upper bounds obtained are respectively given by 18.9711 and 19.1635. The exact average population level obtained from Monte Carlo simulations of the process is 19.1495.

Although we have discussed the bounds only on $\langle \boldsymbol{x} \rangle$ thus far, this approach yields bounds on all moments up to order M. These bounds can be straightforwardly used to infer the bounds on other statistical quantities of interest, e.g., coefficient of variation, skewness, etc. In particular, the coefficient of variation is equal to $\left(\langle \boldsymbol{x}^2 \rangle / \langle \boldsymbol{x} \rangle^2 \right) - 1$. Thus, a lower (upper) bound on the coefficient of variation can be computed by using

the lower (upper) bound of $\langle \boldsymbol{x}^2 \rangle$ and the upper (lower) bound of $\langle \boldsymbol{x} \rangle$.

To illustrate this point, we compute the bounds on the coefficient of variation squared as k/r is varied. It is worth noting that both the upper and lower bounds on $\langle \boldsymbol{x} \rangle$ in (2.25) depend only on k/r, and not on individual values of k and r. This holds true for all bounds obtained here as evident from the steady-state of the moment equation in (2.45). We call the ratio k/r as the relative immigration rate and show its effect on the coefficient of variation in Fig. 2.4 (Right). It is seen that both the upper and lower bounds on coefficient of variation decrease with increase in k/r, thus suggesting that the coefficient of variation decreases. Interestingly, the difference between the bounds is large for small values of k/r and it becomes negligible for high values of k/r. Thus, depending upon a parameter regime, a lower or higher order truncation might be used to obtain bounds within desired accuracy.

2.2.3 Stochastic Gene Expression with Negative Auto–regulation

Consider stochastic expression of an auto-regulating gene represented by the following reactions

$$Gene_{OFF} \underset{k_{off} x_2}{\overset{k_{on}}{\rightleftharpoons}} Gene_{ON}, \quad Gene_{ON} \xrightarrow{k_p} \boldsymbol{B} \times Protein \xrightarrow{\gamma_p} \emptyset. \tag{2.26}$$

Here the gene is assumed to reside in one of the two states: ON (active), and OFF (inactive). The protein is producted at a rate k_p from the ON state whereas there is no protein production when the gene is in the OFF state. Each production event produces B number of protein molecules where B follows a geometric distribution. The gene state is represented by $\boldsymbol{x}_1$, which is a Bernoulli random variable ($\boldsymbol{x}_1 = 1(0)$ for the ON (OFF) state), and the protein level is represented by $\boldsymbol{x}_2$. The gene negatively regulates itself by switching to OFF state in a protein copy number dependent fashion with a rate $k_{off}\boldsymbol{x}_2$. Finally, the protein molecules can degrade with a rate γ_p. This gene expression model has been studied previously [2, 63, 129, 130]; exact solution to its moments are available which allows us to validate the bounds obtained using our method.

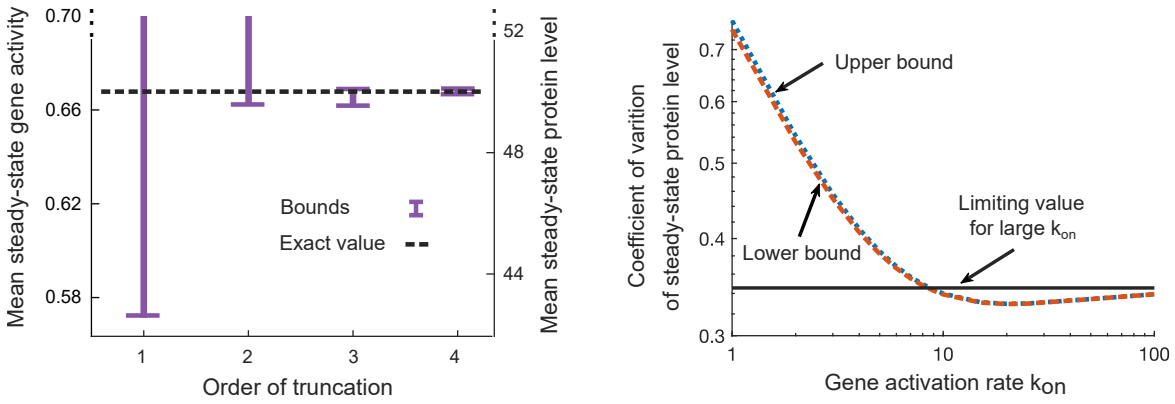

Figure 2.2: Estimated bounds on moments for a stochastic gene expression model. *Left:* The lower and upper bounds for the average gene activity (left axis)/protein level (right axis) for different orders of truncation M is shown. The bound for $M = 1$ correspond to (2.31). As M is increased, the bounds obtained get tighter. The exact mean values are obtained from the analytical solution of the system from [2]. Parameters in (2.26) taken as $k_{on} = 10$, $k_{off} = 0.1$, $k_p = 15$, $\langle \boldsymbol{B} \rangle = 5$, and $\gamma_p = 1$. *Right:* The coefficient of variation (standard deviation/mean) of the steady-state protein level as a function of the gene activation rate k_{on} is plotted. The steady-state protein level in the deterministic sense is kept constant at 50 molecules by varying k_p with k_{on} such that $k_p = 50\gamma_p \left(50k_{off} + k_{on}\right) / \left(k_{on} \langle \boldsymbol{B} \rangle\right)$. The lower and upper bounds on the coefficient of variation are obtained for $M = 5$ and they exhibit U-shape profiles, thus showing that the noise is minimizing at a specific value of k_{on}. Other parameters are taken as $k_{off} = 0.1$, $\langle \boldsymbol{B} \rangle = 5$, and $\gamma_p = 1$.

Table 2.2: Description of reactions for the auto–regulating gene

.

Reaction, $\mathcal{R}_i$	Stoichiometric vector, η_i	Propensity function, $f_i(\boldsymbol{x})$
$Gene_{OFF} \xrightarrow{k_{on}} Gene_{ON}$	$\begin{bmatrix}1 & 0\end{bmatrix}^\top$	$k_{on}(1-\boldsymbol{x}_1)$
$Gene_{ON} \xrightarrow{k_{off_A}\boldsymbol{x}_2} Gene_{OFF}$	$\begin{bmatrix}-1 & 0\end{bmatrix}^\top$	$k_{off}\boldsymbol{x}_1\boldsymbol{x}_2$
$Gene_{ON} \xrightarrow{k_p} Gene_{ON} + \boldsymbol{B} \times Protein$	$\begin{bmatrix}0 & \boldsymbol{B}\end{bmatrix}^\top$	$k_p\boldsymbol{x}_1$
$Protein \xrightarrow{\gamma_p} \emptyset$	$\begin{bmatrix}0 & -1\end{bmatrix}^\top$	$\gamma\boldsymbol{x}_2$

In this example, the state of the system is $\boldsymbol{x} = \begin{bmatrix}\boldsymbol{x}_1 & \boldsymbol{x}_2\end{bmatrix}^\top$. There are four reactions whose propensity functions and stoichiometric vectors are described in Table 2.2. Using (2.2), the time evolution of a moment $\langle \boldsymbol{x}_1^{m_1}\boldsymbol{x}_2^{m_2}\rangle$ is given as

$$\frac{d\langle \boldsymbol{x}_1^{m_1}\boldsymbol{x}_2^{m_2}\rangle}{dt} = \langle k_{on}(1-\boldsymbol{x}_1)\left((\boldsymbol{x}_1+1)^{m_1}-\boldsymbol{x}_1^{m_1}\right)\boldsymbol{x}_2^{m_2} + k_{off}\boldsymbol{x}_1\boldsymbol{x}_2\left((\boldsymbol{x}_1-1)^{m_1}-\boldsymbol{x}_1^{m_1}\right)\boldsymbol{x}_2^{m_2} + k_p\boldsymbol{x}_1^{m_1}\left((\boldsymbol{x}_2+\boldsymbol{B})^{m_2}-\boldsymbol{x}_2^{m_2}\right) + \gamma_p\boldsymbol{x}_2\boldsymbol{x}_1^{m_1}\left((\boldsymbol{x}_2-1)^{m_2}-\boldsymbol{x}_2^{m_2}\right)\rangle. \quad (2.27)$$

Note that dynamics is not closed due to the nonlinearity arising from the negative feedback: a moment $\langle \boldsymbol{x}_1^{m_1}\boldsymbol{x}_2^{m_2}\rangle$ depends upon $\langle \boldsymbol{x}_1^{m_1}\boldsymbol{x}_2^{m_2+1}\rangle$. Furthermore, $\boldsymbol{x}_1 \in \{0,1\}$ is a binary random variable for which the following relations hold

$$\langle \boldsymbol{x}_1^{m_1}\boldsymbol{x}_2^{m_2}\rangle = \langle \boldsymbol{x}_1\boldsymbol{x}_2^{m_2}\rangle, m_1 \in \{1,2,3,\ldots\}, m_2 \in \{0,1,2,3,\ldots\}. \quad (2.28)$$

The above relations imply that the moment vector $\mathcal{X}$ does not need to contain all cross moments. For example, the moments up to order 3 can be stacked as $\mathcal{X} = \begin{bmatrix}1 & \langle \boldsymbol{x}_1\rangle & \langle \boldsymbol{x}_2\rangle & \langle \boldsymbol{x}_1\boldsymbol{x}_2\rangle & \langle \boldsymbol{x}_2^2\rangle & \langle \boldsymbol{x}_1\boldsymbol{x}_2^2\rangle & \langle \boldsymbol{x}_2^3\rangle\end{bmatrix}^\top$. In this case, the corresponding $\overline{\mathcal{X}}$ is given by $\overline{\mathcal{X}} = [\langle \boldsymbol{x}_1\boldsymbol{x}_2^4\rangle]$.

As with the one-dimensional example, here too we are interested in obtaining bounds on moments of the state $\boldsymbol{x}$. Towards this end, we begin by writing the first order moment equations in steady-state

$$\frac{d\langle \boldsymbol{x}_1\rangle}{dt} = k_{on} - k_{on}\langle \boldsymbol{x}_1\rangle - k_{off}\langle \boldsymbol{x}_1\boldsymbol{x}_2\rangle = 0, \quad (2.29)$$

$$\frac{d\langle \boldsymbol{x}_2\rangle}{dt} = k_p\langle \boldsymbol{B}\rangle\langle \boldsymbol{x}_1\rangle - \gamma_p\langle \boldsymbol{x}_2\rangle = 0. \quad (2.30)$$

To obtain a bound on $\langle \boldsymbol{x}_1 \rangle$ and $\langle \boldsymbol{x}_2 \rangle$, we require a bound on the second order moment $\langle \boldsymbol{x}_1 \boldsymbol{x}_2 \rangle$. Generally, the only bound that we can use is $\langle \boldsymbol{x}_1 \boldsymbol{x}_2 \rangle \geq 0$. However, because $\boldsymbol{x}_1$ is a binary random variable, we have that $\langle \boldsymbol{x}_1^2 \boldsymbol{x}_2 \rangle = \langle \boldsymbol{x}_1 \boldsymbol{x}_2 \rangle$. Thus, using the inequality obtained from first matrix of (2.13), a bound $\langle \boldsymbol{x}_1 \boldsymbol{x}_2 \rangle \leq \langle \boldsymbol{x}_2 \rangle$ can be found. Plugging this in the moment equations yields

$$\frac{k_{on}\gamma_p}{k_{on}\gamma_p + k_{off}k_p \langle \boldsymbol{B} \rangle} \leq \langle \boldsymbol{x}_1 \rangle \leq 1, \quad \frac{k_p \langle \boldsymbol{B} \rangle}{\gamma_p} \frac{k_{on}\gamma_p}{k_{on}\gamma_p + k_{off}k_p \langle \boldsymbol{B} \rangle} \leq \langle \boldsymbol{x}_2 \rangle \leq \frac{k_p \langle \boldsymbol{B} \rangle}{\gamma_p}. \tag{2.31}$$

In the similar fashion as above, we can write moment equations up to order two and use the inequality for $\langle \boldsymbol{x}_1 \boldsymbol{x}_2^2 \rangle$ obtained from the second matrix of (2.13). This eventually leads to an improvement in the lower bounds on both $\langle \boldsymbol{x}_1 \rangle$ and $\langle \boldsymbol{x}_2 \rangle$.

Continuing in similar way, we obtain improvements in the lower bounds for $M = 4$ and $M = 6$, and improvements in the upper bounds for $M = 3$, $M = 5$ and $M = 7$. The bounds up to $M = 4$ are shown in Fig. 2.2 (Left). The lower and upper bounds obtained for $\langle \boldsymbol{x}_1 \rangle$ via 7^{th} order truncation are 0.667463 and 0.667465 respectively. These are quite precise as the exact solution for $\langle \boldsymbol{x}_1 \rangle$ is 0.667464 as obtained using the exact solution from [2]. It is worth noting that as discussed below (2.6), we can also use additional bounds arising from the fact that $\boldsymbol{x}_1 \in [0, 1]$, so $\boldsymbol{x}_1$ and $1 - \boldsymbol{x}_1$ both would be positive. However, in this particular example, they do not lead to significant improvements in the bounds.

As done in the logistic growth example, we also obtain the bounds on the second order moment and compute the bounds on coefficient of variation. In particular, we study the effect of varying the parameter k_{on} (gene activation rate). Our results show that there is a U-shape curve, and the noise is minimized at an optimal value of k_{on} (Fig. 2.2 (Right)). Furthermore, the coefficient of variation approaches a limiting value for large values of k_{on} (i.e., the gene is always ON). Similar results were obtained in [63] for an auto-regulating gene expression model using the moment closure techniques. Notably, here the lower and upper bounds remain reasonably close to each other for the range of k_{on} considered as opposed to the logistic growth example.

2.2.4 Activator-Repressor Gene Motif

Next we apply the proposed method to estimate bounds on moments of a slightly more complicated example of a gene network motif that consists of two genes as described by the following reactions:

$$Gene_{A_{OFF}} \underset{k_{off_A} x_4}{\overset{k_{on_A}}{\rightleftharpoons}} Gene_{A_{ON}}, \quad Gene_{A_{ON}} \xrightarrow{k_A} \boldsymbol{B}_A \times Protein_A \xrightarrow{\gamma_A} \emptyset \tag{2.32}$$

$$Gene_{R_{OFF}} \underset{k_{off_R}}{\overset{k_{on_R}\, \boldsymbol{x}_2}{\rightleftharpoons}} Gene_{R_{ON}}, \quad Gene_{R_{ON}} \xrightarrow{k_R} \boldsymbol{B}_R \times Protein_R \xrightarrow{\gamma_R} \emptyset \tag{2.33}$$

Here the notations with subscript A denote the activator whereas those with subscript R represent the repressor. The state of the system is represented by the vector $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 & \boldsymbol{x}_4 \end{bmatrix}^\top$, where $\boldsymbol{x}_1$ and $\boldsymbol{x}_3$ respectively represent the activator and repressor gene state. The corresponding proteins are represented by $\boldsymbol{x}_2$ and $\boldsymbol{x}_4$. As with the previous example, the genes are assumed to reside in two states: OFF and ON. The activator turns the repressor gene ON, whereas, the repressor turns the activator gene OFF, thereby creating a feedback loop. There are eight reactions in this example. The corresponding stoichiometric vectors and propensity functions are given in Table 2.3.

Table 2.3: Description of reactions for the activator repressor motif

Reaction, $\mathcal{R}_i$	Stoichiometric vector, η_i	Propensity function, $f_i(\boldsymbol{x})$
$Gene_{A_{OFF}} \xrightarrow{k_{on_A}} Gene_{A_{ON}}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^\top$	$k_{on_A}(1 - \boldsymbol{x}_1)$
$Gene_{A_{ON}} \xrightarrow{k_{off_A}\boldsymbol{x}_4} Gene_{A_{OFF}}$	$\begin{bmatrix} -1 & 0 & 0 & 0 \end{bmatrix}^\top$	$k_{off_A}\boldsymbol{x}_1\boldsymbol{x}_4$
$Gene_{A_{ON}} \xrightarrow{k_A} Gene_{A_{ON}} + \boldsymbol{B}_A \times Protein_A$	$\begin{bmatrix} 0 & \boldsymbol{B}_A & 0 & 0 \end{bmatrix}^\top$	$k_A\boldsymbol{x}_1$
$Protein_A \xrightarrow{\gamma_A} \emptyset$	$\begin{bmatrix} 0 & -1 & 0 & 0 \end{bmatrix}^\top$	$\gamma_A\boldsymbol{x}_2$
$Gene_{R_{OFF}} \xrightarrow{k_{on_R}\boldsymbol{x}_2} Gene_{R_{ON}}$	$\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^\top$	$k_{on_R}\boldsymbol{x}_2(1 - \boldsymbol{x}_3)$
$Gene_{R_{ON}} \xrightarrow{k_{off_R}} Gene_{R_{OFF}}$	$\begin{bmatrix} 0 & 0 & -1 & 0 \end{bmatrix}^\top$	$k_{off_R}\boldsymbol{x}_3$
$Gene_{R_{ON}} \xrightarrow{k_R} Gene_{R_{ON}} + \boldsymbol{B}_R \times Protein_R$	$\begin{bmatrix} 0 & 0 & 0 & \boldsymbol{B}_R \end{bmatrix}^\top$	$k_R\boldsymbol{x}_3$
$Protein_R \xrightarrow{\gamma_R} \emptyset$	$\begin{bmatrix} 0 & 0 & 0 & -1 \end{bmatrix}^\top$	$\gamma_R\boldsymbol{x}_4$

The time evolution of a moment $\langle \boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \boldsymbol{x}_3^{m_3} \boldsymbol{x}_4^{m_4} \rangle$ can be computed using (2.2). For example, the first order moment equations in the steady state are given by:

$$\frac{d\langle \boldsymbol{x}_1 \rangle}{dt} = k_{on_A} - k_{on_A} \langle \boldsymbol{x}_1 \rangle - k_{off_A} \langle \boldsymbol{x}_1 \boldsymbol{x}_4 \rangle = 0, \tag{2.34}$$

$$\frac{d\langle \boldsymbol{x}_2 \rangle}{dt} = k_A \langle \boldsymbol{B}_A \rangle \langle \boldsymbol{x}_1 \rangle - \gamma_A \langle \boldsymbol{x}_2 \rangle = 0, \tag{2.35}$$

$$\frac{d\langle \boldsymbol{x}_3 \rangle}{dt} = k_{on_R} \langle \boldsymbol{x}_2 \rangle - k_{on_R} \langle \boldsymbol{x}_2 \boldsymbol{x}_3 \rangle - k_{off_R} \langle \boldsymbol{x}_3 \rangle = 0, \tag{2.36}$$

$$\frac{d\langle \boldsymbol{x}_4 \rangle}{dt} = k_R \langle \boldsymbol{B}_R \rangle \langle \boldsymbol{x}_3 \rangle - \gamma_R \langle \boldsymbol{x}_4 \rangle = 0. \tag{2.37}$$

These equations are not closed as the first order moments depend upon the second order moments $\langle \boldsymbol{x}_1 \boldsymbol{x}_4 \rangle$ and $\langle \boldsymbol{x}_2 \boldsymbol{x}_3 \rangle$. Solving the first order moments gives

$$\langle \boldsymbol{x}_1 \rangle = 1 - \frac{k_{off_A}}{k_{on_A}} \langle \boldsymbol{x}_1 \boldsymbol{x}_4 \rangle, \tag{2.38}$$

$$\langle \boldsymbol{x}_2 \rangle = \frac{k_A \langle \boldsymbol{B}_A \rangle}{\gamma_A} \left(1 - \frac{k_{off_A}}{k_{on_A}} \langle \boldsymbol{x}_1 \boldsymbol{x}_4 \rangle \right) \tag{2.39}$$

$$\langle \boldsymbol{x}_3 \rangle = \frac{k_{on_R}}{k_{off_R}} \left(\frac{k_A \langle \boldsymbol{B}_A \rangle}{\gamma_A} \left(1 - \frac{k_{off_A}}{k_{on_A}} \langle \boldsymbol{x}_1 \boldsymbol{x}_4 \rangle \right) - \langle \boldsymbol{x}_2 \boldsymbol{x}_3 \rangle \right), \tag{2.40}$$

$$\langle \boldsymbol{x}_4 \rangle = \frac{k_R \langle \boldsymbol{B}_R \rangle}{\gamma_R} \frac{k_{on_R}}{k_{off_R}} \left(\frac{k_A \langle \boldsymbol{B}_A \rangle}{\gamma_A} \left(1 - \frac{k_{off_A}}{k_{on_A}} \langle \boldsymbol{x}_1 \boldsymbol{x}_4 \rangle \right) - \langle \boldsymbol{x}_2 \boldsymbol{x}_3 \rangle \right). \tag{2.41}$$

Using the property mentioned in (2.28) of Bernoulli random variables $\boldsymbol{x}_1$ and $\boldsymbol{x}_3$, we have that $0 \leq \langle \boldsymbol{x}_1 \boldsymbol{x}_4 \rangle \leq \langle \boldsymbol{x}_4 \rangle$ and $0 \leq \langle \boldsymbol{x}_2 \boldsymbol{x}_3 \rangle \leq \langle \boldsymbol{x}_2 \rangle$. Applying these inequalities in the moment equations yields the trivial bounds, such as $0 \leq \langle \boldsymbol{x}_1 \rangle \leq 1$. As the order of truncation is increased, the number of moment equations and corresponding inequalities grows significantly such that the bounds cannot be obtained with off-the-shelf tools. We therefore employ the semidefinite programming based optimization to obtain lower and upper bounds on moments of interest. As expected, these bounds improve as more moment equations and subsequently semidefinite matrices with higher order moments are used (Fig. 2.3). For $M = 7$, the bounds for both mean and second order moments of the activator and repressor are fairly close to each other.

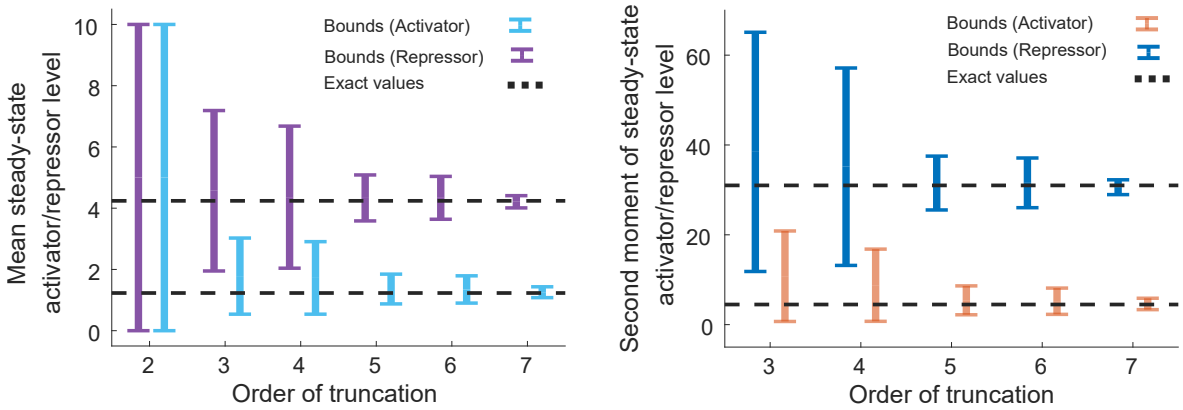

Figure 2.3: Estimated bounds on steady-state moments for the activator repressor motif. *Left:* Upper and lower bounds on the mean of both the activator and repressor are shown for different orders of truncation M. As M is increased, the bounds obtained get tighter. The exact mean values of 1.23 for the activator and 4.24 for the repressor are obtained by averaging $30,000$ MC simulations performed using SSA [1]. Parameters are taken as $k_{on_A} = k_{on_R} = 1$, $k_{off_A} = 5$, $k_{off_R} = 1$, $k_A = k_R = 10$, $\gamma_A = \gamma_R = 1$, and $\boldsymbol{B}_A = \boldsymbol{B}_R = 1$ with probability one. *Right:* The bounds on the second order moments of the steady-state activator/repressor levels are shown for different orders of truncation. The bounds improve with increase in the order of truncation. The exact second order moments of 4.48 for activator and 30.98 for the repressor are shown as obtained from $30,000$ MC simulations. The parameters are taken to be same as those for the left part.

2.3 Bounds on Transient Moments

The method to obtain bounds on stationary moments can also be expanded to moments at a given time point. In this case, the methodology based on inequalities on moments is hard to solve analytically, and one needs to adopt the specialized approach based on semidefinite programming. To this end, the semidefinite program in (2.14) that gives a lower bound on a moment of interest μ at a given time τ can be modified as

$$\begin{aligned}
\underset{\mathcal{X}(t),\overline{\mathcal{X}}(t)}{\text{minimize}} \quad & \mu(\tau) && \text{(2.42a)}\\
\text{subject to} \quad & \frac{d\mathcal{X}}{dt} = A\mathcal{X}(t) + \bar{A}\overline{\mathcal{X}}(t) && \text{(2.42b)}\\
& H_e(t) \succeq 0 && \text{(2.42c)}\\
& H_{o,e}(t) \succeq 0 && \text{(2.42d)}\\
& \mathcal{X}(0) = \mathcal{X}_0 && \text{(2.42e)}
\end{aligned}$$

for all $t \in [0, \tau]$. The upper bound can be computed by maximizing the objective function. The key difference with (2.14) is that (2.42) involves a dynamic constraint imposed by moment dynamics. This constraint can be handled by discretizing the differential equation using Euler's method, and imposing the positive semidefiniteness of $H_{o,i}$ and H_e at all discrete time points.

2.3.1 Stochastic Logistic Model

Consider a version of the stochastic logistic model studied in [59]. This model is described as

$$X \xrightarrow{k_1 - k_2\boldsymbol{x}} X + 1, \quad X \xrightarrow{k_3 + k_4\boldsymbol{x}} X - 1 \tag{2.43}$$

Here we assume that the coefficients satisfy

$$\Omega := \frac{k_1}{k_2} \in \mathbb{N}, \quad k_1 > 0, \quad k_2 > 0, \quad k_3 > 0, \quad k_4 \geq 0. \tag{2.44}$$

These assumptions guarantee that if $\boldsymbol{x}(0) \in \{0, 1, 2, \ldots, \Omega\}$ then $\boldsymbol{x}(t) \in \{0, 1, 2, \ldots, \Omega\}$ for all $t \geq 0$. In this example, there are two reactions with propensity functions and stoichiometric vectors tabulated in Table 2.4.

Table 2.4: Description of reactions for the logistic growth model

Reaction, $\mathcal{R}_i$	Stoichiometric vector, η_i	Propensity function, $f_i(\boldsymbol{x})$
$X \xrightarrow{k_1 - k_2\boldsymbol{x}} X + 1$	$\begin{bmatrix}1\end{bmatrix}^\top$	$k_1\boldsymbol{x} - k_2\boldsymbol{x}^2$
$X \xrightarrow{k_3 + k_4\boldsymbol{x}} X - 1$	$\begin{bmatrix}-1\end{bmatrix}^\top$	$k_3\boldsymbol{x} + k_4\boldsymbol{x}^2$

Using (2.2), we can write the moment dynamics for a m^{th} order moment as

$$\frac{d\left\langle \boldsymbol{x}^m \right\rangle}{dt} = \left\langle \left(k_1\boldsymbol{x} - k_2\boldsymbol{x}^2\right)\left((\boldsymbol{x}+1)^m - \boldsymbol{x}^m\right) + \left(k_3\boldsymbol{x} + k_4\boldsymbol{x}^2\right)\left((\boldsymbol{x}-1)^m - \boldsymbol{x}^m\right)\right\rangle. \quad (2.45)$$

Note that these equations are not closed because of the quadratic propensity functions, and time evolution of a m^{th} order moment depends upon a moment of $(m+1)^{th}$ order. Thus, if we construct a vector $\mathcal{X}$ consisting of first M moments, then $\overline{\mathcal{X}} = [\left\langle \boldsymbol{x}^{M+1} \right\rangle]$.

As discussed above, if $\boldsymbol{x}(0)$ is an integer between 0 and $\Omega = \frac{a_1}{b_1}$, then $\boldsymbol{x}(t)$ is an integer in $[0, \Omega]$ for all time [59]. Recalling the discussion around (2.6), we also utilize bounding constraints $\left\langle \boldsymbol{x}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle \succeq 0$ and $\left\langle (\Omega - \boldsymbol{x})\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top \right\rangle \succeq 0$ at all times. Using the semidefinite program in (2.42), we compute upper and lower bounds on the second moment at the final time, $\langle \boldsymbol{x}(\tau)^2 \rangle$. Surprisingly, the bounds give a good approximation to $\langle \boldsymbol{x}(t)^2 \rangle$ for all $t \in [0, \tau]$, despite only optimizing the bound at the final time, τ.

While the above simple example illustrates the methodology, solving the semidefinite program in (2.42) has several challenges. First, the semidefinite program needs discretization of the time in the interval $[0, \tau]$ and thereby the size of the overall program gets large quickly. Secondly, the semidefinite matrices H_e and H_{o_i} are often ill-conditioned because their elements are moments. Due to these issues, the semidefinite program based approach is computationally restrictive. The program is of course much simpler if bounds on only stationary moments are desired.

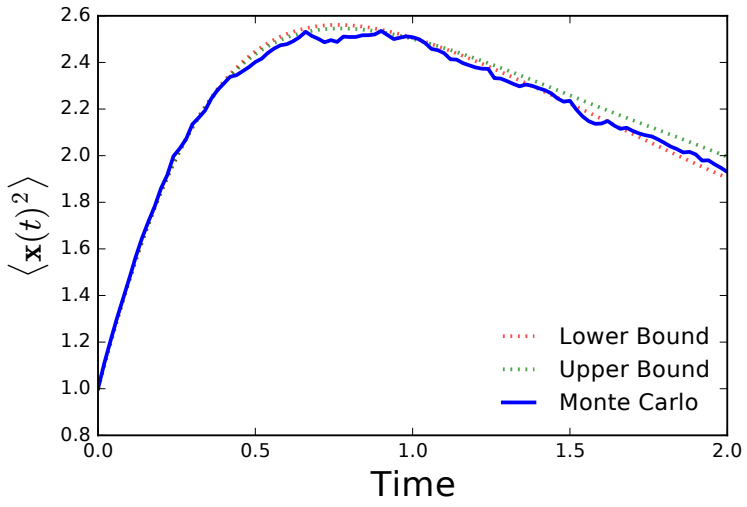

Figure 2.4: Lower and upper bounds on the second moment of the logistic model. For the model in (2.43), maximizing and minimizing the bounds on $\langle \boldsymbol{x}(\tau)^2 \rangle$, we get upper and lower bounds on the true value. The result of 5000 runs of the model are also shown. The parameters used are $k_1 = 3$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0$ and initial condition $\boldsymbol{x}(0) = 1$. While we only penalized the final value of $\langle \boldsymbol{x}(\tau)^2 \rangle$, the upper and lower bound trajectories have similar values over most of the time horizon. Furthermore, these trajectories both give good approximations to the average trajectory found by Monte Carlo simulations.

2.4 Conclusion

We proposed a method to compute lower and upper bounds on moments of a stochastic biochemical reaction system. Our results show that not only one can obtain upper and lower bounds on a given moment, but also both upper and lower bounds improve considerably as one uses more moment equations. Thus, there is a trade-off between the computational cost and the accuracy of bounds.

Chapter 3

BOUNDING MOMENTS OF STOCHASTIC HYBRID SYSTEMS

In the previous chapter, we discussed a method to bound moments of stochastic biochemical systems. Here we extend the framework to Stochastic Hybrid Systems (SHSs). An SHS is a mathematical framework that is applicable to a wide-array of phenomena in engineering, biological and physical systems [131–143]. The biochemical systems described previously are nothing but a sub-class of SHS [119]. Specifying an SHS involves a finite number of discrete states (modes), stochastic dynamics of a continuous state, a set of rules governing transitions that can change the continuous state as well as the discrete state, and reset maps that define how the states change after a transition [133, 144–146].

As with the biochemical systems in the previous chapter, formal analysis of SHSs is often challenging. That is the probability density function of the SHS state space can be characterized by Kolmogorov equations, but solving them analytically is not possible in most cases. Computing moments of SHS is another approach that provides important insights into its dynamics and it turns out that the form of moment dynamics for biochemical systems also holds for some SHSs. In particular, for an SHS whose continuous state, transition intensities, and reset maps are described via polynomials, the time evolution of its moments is governed by a system of linear ordinary differential equations [133]. As one could expect, the issues with moment dynamics of biochemical systems also carry over to the moment dynamics of these polynomial SHSs: it is not closed except for a few special cases [147, 148].

A natural question that arises at this point is whether the methodology to estimate moments in the previous chapter extends to polynomial SHSs. In this chapter, we show that this indeed can be done. Further, in some cases, the SHS might consist

of non-polynomial nonlinearities. For example, in modeling of stochastic biochemical systems, some reaction propensities consist of rational functions [119]. Here, the moment dynamics also contains non-polynomial moments, in addition to the higher order moments as in the polynomial case. We cast the SHS with rational functions as polynomial SHSs by defining additional states [149, 150]. By doing so, we obtain approximate values of desired moments with provable guarantees.

3.1 Background on Stochastic Hybrid Systems

In this section, we provide brief overview of a SHS construction and its mathematical characterization. The reader is referred to [133, 144–146] for technical details on SHS, and its relationship with various other classes of stochastic systems.

3.1.1 Basic Setup

The state space of a SHS consists of a continuous state $\boldsymbol{x}(t) \in \mathbb{R}^n$ and a discrete state $\boldsymbol{q}(t) \in Q = \{s_1, s_2, \ldots, s_N\}$. There are three components of SHS that define how its states evolve over time. First, the continuous state evolves as per a stochastic differential equation (SDE)

$$d\boldsymbol{x} = f(\boldsymbol{q}, \boldsymbol{x})dt + g(\boldsymbol{q}, \boldsymbol{x})d\boldsymbol{w}, \tag{3.1a}$$

where $f : Q \times \mathbb{R}^n \to \mathbb{R}^n$ and $g : Q \times \mathbb{R}^n \to \mathbb{R}^{n\times k}$ are respectively the drift and diffusion terms, and $\boldsymbol{w}$ is a k–dimensional Weiner process. Second, the state $(\boldsymbol{q}, \boldsymbol{x})$ changes stochastically through S transitions/resets that are characterized by the transition intensities

$$\lambda_r(\boldsymbol{q}, \boldsymbol{x}), \quad \lambda_r : Q \times \mathbb{R}^n \to [0, \infty), \quad r = 1, 2, \ldots, S. \tag{3.1b}$$

Third, the transition for each $r = 1, 2, \ldots, S$ has an associated reset map

$$\begin{aligned} &(\boldsymbol{q}, \boldsymbol{x}) \mapsto (\theta_r(\boldsymbol{q}), \phi_r(\boldsymbol{q}, \boldsymbol{x})), \\ &\theta_r : Q \to Q, \phi_r : Q \times \mathbb{R}^n \to \mathbb{R}^n \end{aligned} \tag{3.1c}$$

that defines how the pre-transition discrete and continuous states map into the post-transition discrete and continuous states. One way to think about an SHS is to consider the discrete states as different modes, each of which has an associated SDE describing the time evolution of the continuous state. The reset events can either reset the continuous state and remain in the same mode (i.e, the continuous state evolves via the same SDE as before the reset occured), or reset both the continuous state and the mode.

For purpose of this work, we first assume that for a given discrete state, the functions f, g, λ_r, and ϕ_r are polynomials in $\boldsymbol{x}$. We then consider the case when these could be non-polynomial functions that are composition of rational functions, trigonometric functions, exponential, and logarithm.

3.1.2 Extended Generator

Mathematical characterization of SHS (3.1) requires computation of expectation of some large class of functions evaluated on its state space. To this end, the extended generator describes time evolution of a scalar test function $\psi : Q \times \mathbb{R}^n \to \mathbb{R}$ which is twice continuously differentiable with respect to its second argument (i.e., $\boldsymbol{x}$). This is given as

$$\frac{d\mathbb{E}\left[\psi(\boldsymbol{q},\boldsymbol{x})\right]}{dt} = \mathbb{E}\left[(\mathcal{L}\psi)(\boldsymbol{q},\boldsymbol{x})\right], \tag{3.2a}$$

where $\mathbb{E}$ denotes the expectation operator and $\mathcal{L}$ is called the extended generator

$$\begin{aligned}(\mathcal{L}\psi)(\boldsymbol{q},\boldsymbol{x}) := \frac{\partial\psi(\boldsymbol{q},\boldsymbol{x})}{\partial\boldsymbol{x}} f(\boldsymbol{q},\boldsymbol{x}) \\ + \frac{1}{2}\,\mathrm{Trace}\left(\frac{\partial^2\psi(\boldsymbol{q},\boldsymbol{x})}{\partial\boldsymbol{x}^2} g(\boldsymbol{q},\boldsymbol{x}) g(\boldsymbol{q},\boldsymbol{x})^\top\right) \\ + \sum_{r=1}^{S} \left(\psi\left(\theta_r(\boldsymbol{q}), \phi_r(\boldsymbol{q},\boldsymbol{x})\right) - \psi(\boldsymbol{q},\boldsymbol{x})\right) \lambda_r(\boldsymbol{q},\boldsymbol{x}).\end{aligned} \tag{3.2b}$$

The terms $\frac{\partial\psi(\boldsymbol{q},\boldsymbol{x})}{\partial\boldsymbol{x}}$ and $\frac{\partial^2\psi(\boldsymbol{q},\boldsymbol{x})}{\partial\boldsymbol{x}^2}$ respectively denote the gradient and the Hessian of $\psi(\boldsymbol{q},\boldsymbol{x})$ with respect to $\boldsymbol{x}$ [133]. Appropriate choice of $\psi(\boldsymbol{q},\boldsymbol{x})$ gives a dynamics of moments of SHS as described in the next section.

3.2 Moment Analysis of Polynomial SHS

In this section, we focus on SHS defined over polynomials: for each discrete state $\boldsymbol{q}$, the functions f, g, λ_r, and ϕ_r are polynomials in the continuous state $\boldsymbol{x}$. We describe how the extended generator gives time evolution of its moments. We then discuss the problem of moment closure, and propose our methodology to estimate moments.

3.2.1 Moment Dynamics for Polynomial SHS with Single Discrete State

We first consider a simpler system that has only one discrete mode/state ($\boldsymbol{q}$ can be dropped for ease of notation). For a given n-tuple $m = (m_1, m_2, \ldots, m_n) \in \mathbb{N}^n$, moment dynamics can be computed by plugging in the monomial test function

$$\psi(\boldsymbol{x}) = \boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \ldots \boldsymbol{x}_n^{m_n} \tag{3.3}$$

in (3.2). Here order of the moment $\mathbb{E}(\boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \ldots \boldsymbol{x}_n^{m_n})$ is given by $\sum_{i=1}^n m_i$, and there are $\binom{\sum_{i=1}^n m_i + n - 1}{n-1}$ moments of the order of order $\sum_{i=1}^n m_i$. The following standard result shows how dynamics of a collection of moments of $\boldsymbol{x}$ evolves over time for a special class of SHS that are defined via polynomials.

Lemma 1 *Let $f(\boldsymbol{x})$, $g(\boldsymbol{x})g(\boldsymbol{x})^\top$, $\lambda_r(\boldsymbol{x})$ and $\phi_r(\boldsymbol{x})$ be polynomials in $\boldsymbol{x}$. Denoting the vector consisting of all moments up to a specific order of $\boldsymbol{x}$ by $\mathcal{X}$, its time evolution can be compactly written as*

$$\frac{d\mathcal{X}}{dt} = A\mathcal{X} + \bar{A}\,\overline{\mathcal{X}} \tag{3.4}$$

for appropriately defined matrices A, $\bar{A}$. Here $\overline{\mathcal{X}}$ is a collection of moments whose order is higher than those stacked up in $\mathcal{X}$.

Proof: Since $f(\boldsymbol{x})$, $g(\boldsymbol{x})g(\boldsymbol{x})^\top$, $\lambda_r(\boldsymbol{x})$ and $\phi_r(\boldsymbol{x})$ are polynomials, the extended generator in (3.2b) maps monomials of the form $\boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \ldots \boldsymbol{x}_n^{m_n}$ to a linear combination of monomials of different orders. Collecting all moments up to some order (including the zeroth order moment) in a vector $\mathcal{X}$, the form in (3.4) follows from (3.2a). ■

The moment dynamics in (3.4) is well-known [133]. It is worth noting that the matrix $\bar{A}$ has all its elements *zero* if all the functions $f(\boldsymbol{x})$, $g(\boldsymbol{x})g(\boldsymbol{x})^\top$, $\lambda_r(\boldsymbol{x})$ and $\phi_r(\boldsymbol{x})$

are affine in $\boldsymbol{x}$. In this case, the moments contained in $\mathcal{X}$ can be exactly computed. Next, we discuss moment dynamics for SHS with multiple discrete states.

3.2.2 Moment Dynamics for Polynomial SHS with Finite Number of Discrete States

Now we consider a general SHS that has a finite, but more than one, discrete states. In this case, one is interested in knowing moments of the continuous state given a discrete state and the probability that the system is in the given discrete state. To compute these, we define an N-dimensional state

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_N) \in \mathbb{R}^N \tag{3.5a}$$

such that each $\boldsymbol{\beta}_i, i = 1, 2, \dots, N$ serves as an indicator of the discrete state being $\boldsymbol{q} = s_i$

$$\boldsymbol{\beta}_i = \begin{cases} 1, & \boldsymbol{q} = s_i \\ 0, & \text{else} \end{cases}. \tag{3.5b}$$

For example, when the discrete state $\boldsymbol{q} = s_1$, then we represent it by the tuple $\boldsymbol{\beta} = (1, 0, \dots, 0)$. It follows that the following properties hold

$$\sum_{i=1}^{N} \boldsymbol{\beta}_i = 1; \quad \boldsymbol{\beta}_i \boldsymbol{\beta}_j = 0,\ i \neq j; \quad \boldsymbol{\beta}_i^2 = \boldsymbol{\beta}_i. \tag{3.5c}$$

Furthermore, $\mathbb{E}(\boldsymbol{\beta}_i)$ is equal to the probability of $\boldsymbol{q} = s_i$, while $\mathbb{E}(\boldsymbol{\beta}_i \boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \dots \boldsymbol{x}_n^{m_n})$ is equal to the product of the the probability that $\boldsymbol{q} = s_i$ and the moment of $\boldsymbol{x}_1^{m_1} \boldsymbol{x}_2^{m_2} \dots \boldsymbol{x}_n^{m_n}$, conditioned on $\boldsymbol{q} = s_i$. We can recast the SHS in (3.1) to the new state space $(\boldsymbol{\beta}, \boldsymbol{x})$ as described via the following lemma.

Lemma 2 *Consider the SHS described in (3.1). With $\boldsymbol{\beta} \in \mathbb{R}^N$ defined in (3.5), let a single-discrete mode SHS with state space $(\boldsymbol{\beta}, \boldsymbol{x}) \in \mathbb{R}^{N+n}$ be described by the continuous dynamics*

$$d \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \sum_{i=1}^{N} \boldsymbol{\beta}_i f(s_i, \boldsymbol{x}) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \sum_{i=1}^{N} \boldsymbol{\beta}_i g(s_i, \boldsymbol{x}) d\boldsymbol{w} \end{bmatrix}, \tag{3.6a}$$

reset intensities

$$\sum_{i=1}^{N} \boldsymbol{\beta}_i \lambda_r(s_i, \boldsymbol{x}), \quad r = 1, 2, \ldots, S, \tag{3.6b}$$

and reset maps

$$(\boldsymbol{\beta}, \boldsymbol{x}) \mapsto \left(\boldsymbol{\beta} - \sum_{i=1}^{N} \boldsymbol{\beta}_i \mathbb{1}_{s_i} + \sum_{i=1}^{N} \boldsymbol{\beta}_i \mathbb{1}_{\theta_r(s_i)}, \sum_{i=1}^{N} \boldsymbol{\beta}_i \phi_r(s_i, \boldsymbol{x}) \right), \tag{3.6c}$$

where $\mathbb{1}_{s_i}$ denotes an N-dimensional vector consisting of zeros except for i^{th} position. Then (3.6) recasts (3.1) in $(\boldsymbol{\beta}, \boldsymbol{x})$ space.

Proof: Let $\boldsymbol{q}(t) = s_j \in Q$. Then (3.5) implies that dynamics of $\boldsymbol{x}$ in (3.6a) becomes

$$d\boldsymbol{x} = f(s_j, \boldsymbol{x})dt + g(s_j, \boldsymbol{x})d\boldsymbol{w}, \tag{3.7}$$

which is same as (3.1a). Likewise, the rest intensities for both (3.6) and (3.1) take the form

$$\lambda_r(s_j, \boldsymbol{x}), \quad r = 1, 2, \ldots, S. \tag{3.8}$$

As for the reset maps, (3.6c) yields

$$\left(\mathbb{1}_{s_j}, \boldsymbol{x}\right) \mapsto \left(\mathbb{1}_{s_j} - \mathbb{1}_{s_j} + \mathbb{1}_{\theta_r(s_j)}, \phi_r(s_j, \boldsymbol{x})\right), \tag{3.9}$$

which by definition in (3.5) is same as (3.62)

$$(s_j, \boldsymbol{x}) \mapsto (\theta_r(s_j), \phi_r(s_j, \boldsymbol{x})). \tag{3.10}$$

Since we arbitrarily chose $\boldsymbol{q} = s_j \in Q$, the equivalence between the two SHSs will hold true for any $\boldsymbol{q}$. ■

To write the moment dynamics of SHS in (3.6), we can use monomial test functions

$$\psi(\boldsymbol{\beta}, \boldsymbol{x}) = \boldsymbol{\beta}_1^{m_1} \boldsymbol{\beta}_2^{m_2} \ldots \boldsymbol{\beta}_N^{m_N} \boldsymbol{x}_1^{m_{N+1}} \boldsymbol{x}_2^{N+m_2} \ldots \boldsymbol{x}_n^{m_{N+n}}, \tag{3.11}$$

supplemented with the constraints in (3.5c). It is worth noting that (3.6) is a polynomial SHS in $(\boldsymbol{\beta}, \boldsymbol{x})$ space if the original SHS was polynomial in $\boldsymbol{x}$. The following result provides a general form for the moment dynamics.

Theorem 1 *Consider the SHS in (3.6). Let f, g, λ_l and ϕ_l be polynomials in $\boldsymbol{x}$. Denoting the vector consisting of all moments up to a specific order of the state $(\boldsymbol{\beta}, \boldsymbol{x})$ by $\mathcal{X}$, its time evolution can be compactly written as*

$$\frac{d\mathcal{X}}{dt} = A\mathcal{X} + \bar{A}\overline{\mathcal{X}}, \tag{3.12a}$$

$$0 = C\mathcal{X} + \bar{C}\overline{\mathcal{X}} \tag{3.12b}$$

for appropriately defined matrices A, $\bar{A}$, C, $\bar{C}$. Here $\bar{\mathcal{X}}$ is a collection of moments whose order is higher than those stacked up in $\mathcal{X}$.

Proof: Since (3.6) is polynomial in $(\boldsymbol{\beta}, \boldsymbol{x})$, the form in (3.12a) follows from Lemma 1. The property $\boldsymbol{\beta}_i\boldsymbol{\beta}_j = 0$ in (3.5c) implies that for a non-zero $m_i \in \mathbb{N}$, all moments except those of the form $\mathbb{E}\left(\boldsymbol{\beta}_i^{m_i}\boldsymbol{x}_1^{m_{N+1}}\boldsymbol{x}_1^{m_{N+2}}\dots\boldsymbol{x}_n^{m_{N+n}}\right)$ are zero. Furthermore, $\boldsymbol{\beta}_i^2 = \boldsymbol{\beta}_i$ results in

$$\mathbb{E}\left(\boldsymbol{\beta}_i^{m_i}\boldsymbol{x}_1^{m_{N+1}}\boldsymbol{x}_1^{m_{N+2}}\dots\boldsymbol{x}_n^{m_{N+n}}\right) = \mathbb{E}\left(\boldsymbol{\beta}_i\boldsymbol{x}_1^{m_{N+1}}\boldsymbol{x}_1^{m_{N+2}}\dots\boldsymbol{x}_n^{m_{N+n}}\right), \tag{3.13}$$

for all $m_i \geq 1$. The constraint $\sum_{i=1}^{N}\boldsymbol{\beta}_i = 1$ results in

$$\sum_{i=1}^{N}\mathbb{E}\left(\boldsymbol{\beta}_i\boldsymbol{x}_1^{m_{N+1}}\boldsymbol{x}_1^{m_{N+2}}\dots\boldsymbol{x}_n^{m_{N+n}}\right) - \mathbb{E}\left(\boldsymbol{x}_1^{m_{N+1}}\boldsymbol{x}_1^{m_{N+2}}\dots\boldsymbol{x}_n^{m_{N+n}}\right) = 0. \tag{3.14}$$

These three constraints can be compactly represented by (3.12b). ■

In Theorem 1 we have assumed that all moments up to a certain order are collected in $\mathcal{X}$ and remaining, higher order, moments are collected in $\overline{\mathcal{X}}$. However, since many of these moments are equal to zero, in practice we do not include them in $\mathcal{X}$ and $\overline{\mathcal{X}}$. Similarly, higher order moments that are equal to lower order moments, as in (3.13), are not included.

The form of moment dynamics for polynomial SHSs implies that the moments in $\mathcal{X}$ cannot be computed exactly, since they depend upon the moments in $\overline{\mathcal{X}}$. This is often referred to the problem of moment closure, and there are many methods that have been proposed in the literature to close the moment dynamics. Some of these methods ignore the higher order moments or cumulants to find the closure, while others use dynamical systems properties or physical principles to find the closure [149, 151–153].

In all these methods, the approximations are ad-hoc; they could be quite accurate for a specific system under study while they could perform poorly for other systems.

Recall the semidefinite program proposed in (2.42). The moment dynamics in single-mode, polynomial SHS case has exact same form as the biochemical reaction system considered in the previous chapter, and therefore the semidefinite program trivially applies to the SHS setup. For the multimode polynomial SHS, we have an additional linear constraint $C\mathcal{X}(t) + \bar{C}\overline{\mathcal{X}}(t) = 0$ that needs to be appended to (2.42)

$$\begin{aligned}
\underset{\mathcal{X}(t),\overline{\mathcal{X}}(t)}{\text{minimize}} \quad & \mu(\tau) && \text{(3.15a)}\\
\text{subject to} \quad & \frac{d\mathcal{X}}{dt} = A\mathcal{X}(t) + \bar{A}\overline{\mathcal{X}}(t) && \text{(3.15b)}\\
& 0 = C\mathcal{X}(t) + \bar{C}\overline{\mathcal{X}}(t) && \text{(3.15c)}\\
& H_e(t) \succeq 0 && \text{(3.15d)}\\
& H_{o_i}(t) \succeq 0 && \text{(3.15e)}\\
& \mathcal{X}(0) = \mathcal{X}_0 && \text{(3.15f)}
\end{aligned}$$

for all $t \in [0, \tau]$. Next, we discuss how this method could be extended to rational SHS.

3.3 Moment Analysis for Rational SHS

Consider the SHS in (3.1) wherein all functions are rationals except for the reset maps ϕ_l which we assume to be polynomial. Without loss of generality, we can consider a single discrete state since Lemma 2 allows reduction of a SHS with multiple discrete modes. Let $J(\boldsymbol{x})$ be the least common denominator for all $f(\boldsymbol{x})$, $g(\boldsymbol{x})g^\top(\boldsymbol{x})$, and $\lambda_l(\boldsymbol{x})$. Defining a new state $\boldsymbol{y} = \frac{1}{J(\boldsymbol{x})}$, it is straightforward to see that one gets a polynomial SHS in the state $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{n+1}$, with an equality constraint

$$J(\boldsymbol{x})\boldsymbol{y} - 1 = 0. \tag{3.16}$$

While not studied formally in the context of SHSs, a similar approach to define additional states to study non-polynomial stochastic systems has been used earlier [149, 150]. For SHSs with rational functions, if the reset map is rational then each

monomial in $\boldsymbol{x}$ is mapped to a different rational function and a lot many additional states may be required to define moment dynamics up to a certain order. In light of this, the above method may seem bit restrictive, but in practice there are numerous examples of SHSs wherein only polynomial reset maps appear. In the following Theorem, we provide a general form of moment dynamics for non-polynomial SHSs that can be casted as a polynomial SHS with constraints of the form (3.16).

Theorem 2 *Consider a single discrete mode SHS in* (3.1) *with constraints of the form* (3.16)*. Collecting moments of the state space* $(\boldsymbol{x}, \boldsymbol{y})$ *up to a specific order in the vector* $\mathcal{X}$*, the moment dynamics is given by*

$$\frac{d\mathcal{X}}{dt} = A\mathcal{X} + \bar{A}\overline{\mathcal{X}}, \tag{3.17a}$$

$$0 = C_e\mathcal{X} + \bar{C}_e\overline{\mathcal{X}} \tag{3.17b}$$

where $\overline{\mathcal{X}}$ *contains moments of higher order and the matrices* A, $\bar{A}$, C_e, $\bar{C}_e$ *are appropriately defined.*

Proof: Since the test function is monomial of the form $\boldsymbol{x}_1^{m_1}\boldsymbol{x}_2^{m_2}\dots\boldsymbol{x}_n^{m_n}$, these are closed under the extended generator. Thus, (3.17a) follows from Lemma 1. The algebraic constraints of the form $J(\boldsymbol{x})\boldsymbol{y} - 1 = 0$ imply that the moments in which elements of $J(\boldsymbol{x})\boldsymbol{y} - 1$ appear are equal to zero. Similar to (3.12b), these are encoded as (3.17b). ■

We can straightforwardly extend the above form of moment dynamics to an SHS with multiple discrete states by virtue of Lemma 2

3.3.1 Bounds on Moments via Semidefinite Programming

The preceding discussion provides a recipe to write a rational SHS as polynomial SHS with algebraic constraints. Formally the semidefinite program is now given by

$$\begin{aligned}
\underset{\mathcal{X}(t),\overline{\mathcal{X}}(t)}{\text{minimize}} \quad & \mu(\tau) && \text{(3.18a)}\\
\text{subject to} \quad & \frac{d\mathcal{X}}{dt} = A\mathcal{X}(t) + \bar{A}\overline{\mathcal{X}}(t) && \text{(3.18b)}\\
& 0 = C_e\mathcal{X}(t) + \bar{C}_e\overline{\mathcal{X}}(t) && \text{(3.18c)}\\
& H_e(t) \succeq 0 && \text{(3.18d)}\\
& H_{o_i}(t) \succeq 0 && \text{(3.18e)}\\
& \mathcal{X}(0) = \mathcal{X}_0 && \text{(3.18f)}
\end{aligned}$$

for all $t \in [0, \tau]$. As mentioned earlier, if a multimode SHS were to be considered, the form of semidefinite program remains to be similar with another constraint $C\mathcal{X}+\bar{C}\overline{\mathcal{X}} = 0$ being added. Further, in the steady-state version of program, the moment dynamics constraint needs to be equal to zero.

3.4 Numerical Examples of Moment Estimation

We illustrate our approach using three examples. As discussed in the previous chapter, the methodology is more suitable for studying the steady-state moments. Therefore, here we only provide examples of estimating stationary moments of SHSs. The first example is a simple SHS with single discrete mode, the second example comprises of multiple discrete states and polynomial dynamics/resets, and the second example consists of a single discrete state with rational dynamics. The first two examples are taken and slightly modified from [144] wherein these are studied with a moment closure technique.

3.4.1 Network Control System

Consider a scalar linear stochastic system

$$d\boldsymbol{x} = k_1\boldsymbol{x}dt + k_2 d\boldsymbol{w}, \tag{3.19}$$

whose state is being estimated through measurements received over a communication network. Then, the error $\boldsymbol{e}$ in state estimation can be generated by a SHS with continuous dynamics

$$d\boldsymbol{e} = k_1 \boldsymbol{e} dt + k_2 d\boldsymbol{w}. \tag{3.20}$$

The error is reset to zero

$$\boldsymbol{e} \mapsto \phi(\boldsymbol{e}) = 0, \tag{3.21}$$

whenever the measurement is received (see Fig. 3.1). The intensity of the reset, i.e., the transmission rate of the messages is assumed to be dependent on the continuous state e. For purpose of this work, we take the reset intensity to be

$$\lambda = \boldsymbol{e}^2. \tag{3.22}$$

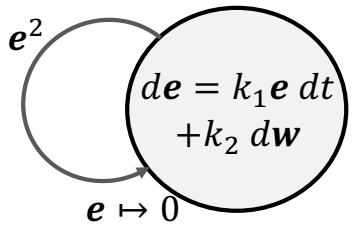

Figure 3.1: Stochastic Hybrid System representation of the error dynamics in a network control system. Here, the system consists of one discrete mode, and the error (continuous dynamics) evolves as per a stochastic differential equation. Upon the reset (which occurs with intensity e^2), the error resets to zero.

Using (3.2b), the extended generator can be written as

$$(L\psi)(\boldsymbol{e}) = k_1 \boldsymbol{e} \frac{\partial \psi(\boldsymbol{e})}{\partial \boldsymbol{e}} + \frac{k_2^2}{2} \frac{\partial^2 \psi(\boldsymbol{e})}{\partial \boldsymbol{e}^2} + \boldsymbol{e}^2 (\psi(0) - \psi(\boldsymbol{e})). \tag{3.23}$$

Let the test function in (3.23) be

$$\psi(\boldsymbol{e}) = \boldsymbol{e}^m, \tag{3.24}$$

then the moment of m^{th} order evolves as

$$\frac{d\mathbb{E}(\boldsymbol{e}^m)}{dt} = k_1 m \mathbb{E}(\boldsymbol{e}^m) + \frac{m(m-1)k_2^2}{2} \mathbb{E}\left(\boldsymbol{e}^{m-2}\right) - \mathbb{E}\left(\boldsymbol{e}^{m+2}\right). \tag{3.25}$$

Thus, if the first m moments (including the zeroth-order moment which is equal to one) are stacked up in $\mathcal{X}$ then

$$\overline{\mathcal{X}} = \begin{bmatrix} \mathbb{E}\left(\boldsymbol{e}^{m+1}\right) \\ \mathbb{E}\left(\boldsymbol{e}^{m+2}\right) \end{bmatrix}. \tag{3.26}$$

For this example, we compute bound on the second order moment $\mathbb{E}\left(\boldsymbol{e}^2\right)$ in stationary state. Solving this semidefinite program for $k_1 = 1$ and $k_2 = 5$ results in the trivial minimum bound 0 and maximum bound 6.99. On increasing the number moment equations and correspondingly the size of M, it is seen that the bounds improve (see Fig. 3.2).

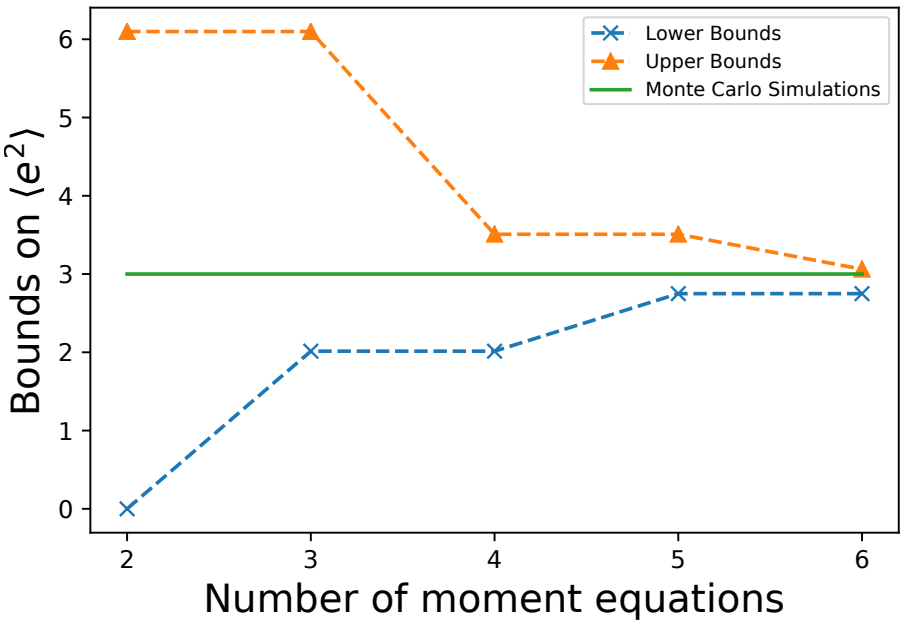

Figure 3.2: Lower and upper bounds on stationary second order moment of the network control system. The parameters are taken to be $k_1 = 1$ and $k_2 = 5$. The bounds converge to the true value as more number of moment equations are used.

3.4.2 TCP On–Off

We consider a simple version of the TCP on-off model. Here, the continuous state of the model is denoted by $\boldsymbol{v}$, which represents the congestion window size of the

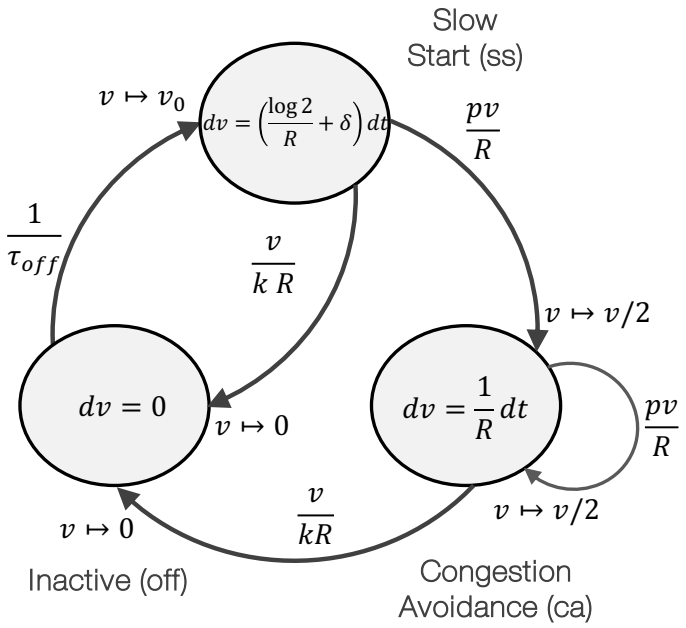

Figure 3.3: Stochastic Hybrid System representation of TCP On-Off model. Here, there are three discrete modes and the continuous dynamics $\boldsymbol{v}$ evolves as per different differential equations depending upon which mode the system is operating in. Various reset intensities and reset maps are also shown.

TCP. The model consists of three discrete states, namely, $\{off, ss, ca\}$, which stand for off, slow start, and congestion avoidance, respectively.

During these modes, the continuous-state evolves as

$$f(\boldsymbol{q}, \boldsymbol{v}, t) = \begin{cases} 0, & \boldsymbol{q} = off \\ \frac{\log 2}{R}\boldsymbol{v} + \delta, & \boldsymbol{q} = ss \\ \frac{1}{R}, & \boldsymbol{q} = ca \end{cases} \tag{3.27}$$

The transitions between the discrete modes are of three types: drop occurences, which correspond to transitions from the ss and ca modes to the ca mode; start of new flow, which correspond to the transitions from the off mode to the ss mode; and termination of flows, which correspond to transitions from the ss and ca modes to the

off mode. These transitions are described via the reset maps

$$\phi_{drop}(\boldsymbol{q},\boldsymbol{v}) = \begin{cases} \left(ca, \frac{\boldsymbol{v}}{2}\right), & \boldsymbol{q} \in \{ss, ca\} \\ (off, \boldsymbol{v}), & \boldsymbol{q} = off \end{cases} \tag{3.28}$$

$$\phi_{start}(\boldsymbol{q},\boldsymbol{v}) = \begin{cases} (\boldsymbol{q}, \boldsymbol{v}), & \boldsymbol{q} \in \{ss, ca\} \\ (ss, v_0), & \boldsymbol{q} = off \end{cases} \tag{3.29}$$

$$\phi_{end}(\boldsymbol{q},\boldsymbol{v}) = \begin{cases} (off, 0), & \boldsymbol{q} \in \{ss, ca\} \\ (off, \boldsymbol{v}), & \boldsymbol{q} = off \end{cases} \tag{3.30}$$

with reset intensities

$$\lambda_{drop}(\boldsymbol{q},\boldsymbol{v}) = \begin{cases} \frac{p\boldsymbol{v}}{R}, & \boldsymbol{q} \in \{ss, ca\} \\ 0, & \boldsymbol{q} = off \end{cases} \tag{3.31}$$

$$\lambda_{start}(\boldsymbol{q},\boldsymbol{v}) = \begin{cases} 0, & \boldsymbol{q} \in \{ss, ca\} \\ \frac{1}{\tau_{off}}, & \boldsymbol{q} = off \end{cases} \tag{3.32}$$

$$\lambda_{end}(\boldsymbol{q},\boldsymbol{v}) = \begin{cases} \frac{\boldsymbol{v}}{kR}, & \boldsymbol{q} \in \{ss, ca\} \\ 0, & \boldsymbol{q} = off \end{cases}. \tag{3.33}$$

Here R is the round trip time, p is the packet drop rate parameter.

To write moment dynamics, we define the indicator state variables $\boldsymbol{\beta}_{ss}$, $\boldsymbol{\beta}_{ca}$, and $\boldsymbol{\beta}_{off}$ as in (3.5b). The resulting single-mode SHS is shown in Fig. 3.4. Using extended generator, we write dynamics of the non-zero moments. In particular, we have

$$\begin{aligned} \frac{d\mathbb{E}(\boldsymbol{\beta}_{ss}\boldsymbol{v}^m)}{dt} &= \frac{m\log 2}{R}\mathbb{E}\left(\boldsymbol{\beta}_{ss}\boldsymbol{v}^m\right) + m\delta\mathbb{E}\left(\boldsymbol{\beta}_{ss}\boldsymbol{v}^{m-1}\right) \\ &\quad + \frac{v_0^m}{\tau_{off}}\mathbb{E}(\boldsymbol{\beta}_{off}) - \left(p + \frac{1}{k}\right)\frac{\mathbb{E}(\boldsymbol{\beta}_{ss}\boldsymbol{v}^{m+1})}{kR}, \end{aligned} \tag{3.34a}$$

$$\frac{d\mathbb{E}(\boldsymbol{\beta}_{ca}\boldsymbol{v}^m)}{dt} = \frac{m}{R}\mathbb{E}\left(\boldsymbol{\beta}_{ca}\boldsymbol{v}^{m-1}\right) + \frac{p}{2^m R}\mathbb{E}(\boldsymbol{\beta}_{ss}\boldsymbol{v}^{m+1}) - \left(\frac{p(2^m-1)}{2^m R} + \frac{1}{kR}\right)\mathbb{E}(\boldsymbol{\beta}_{ca}\boldsymbol{v}^{m+1}), \tag{3.34b}$$

$$\frac{d\mathbb{E}(\boldsymbol{\beta}_{off}\boldsymbol{v}^m)}{dt} = -\frac{\mathbb{E}(\boldsymbol{\beta}_{off}\boldsymbol{v}^m)}{\tau_{off}} + \frac{\mathbb{E}(\boldsymbol{\beta}_{ss}\boldsymbol{v}^{m+1})}{kR} + \frac{\mathbb{E}(\boldsymbol{\beta}_{ca}\boldsymbol{v}^{m+1})}{kR}, \tag{3.34c}$$

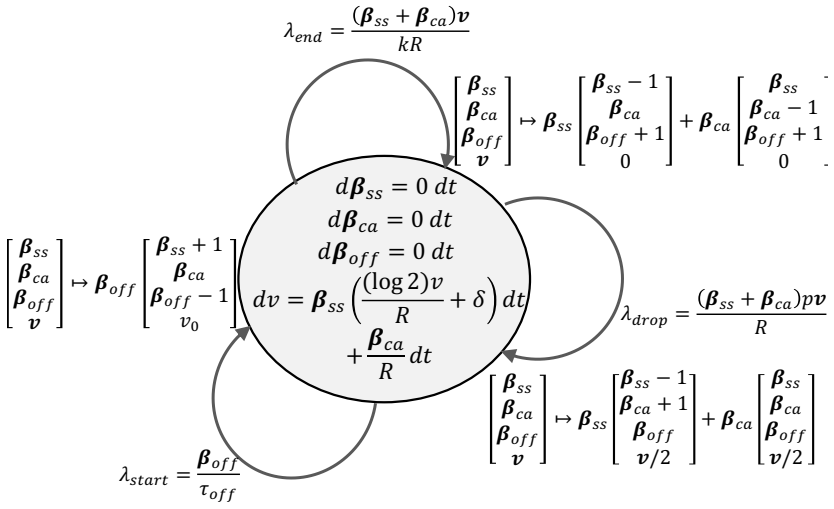

Figure 3.4: An equivalent single-mode representation of the Stochastic Hybrid System representation of TCP On-Off model.

for $m \in \mathbb{N}$. Using these moment equations along with the semidefinite constraints and algebraic constraints arising from the definition of $\boldsymbol{\beta}_{ss}, \boldsymbol{\beta}_{ca}, \boldsymbol{\beta}_{off}$, the semidefinite program can be set up. We can also generate matrices M_i by using non-negativity of $\boldsymbol{\beta}_{ss}, \boldsymbol{\beta}_{ca}, \boldsymbol{\beta}_{off}$, $1 - \boldsymbol{\beta}_{ss}, 1 - \boldsymbol{\beta}_{ca}, 1 - \boldsymbol{\beta}_{off}$. Taking specific values of $R = 5$, $\tau_{off} = 0.5$, $k = 20$, $p = 0.05$, $v_0 = 1$, we get $0.0252 \leq \mathbb{E}(\boldsymbol{\beta}_{ss}) \leq 1$ by utilizing moments of order 2. Considering higher order moments improves these estimates, and we get $0.0912 \leq \mathbb{E}(\boldsymbol{\beta}_{ss}) \leq 0.115$ for moments of order 7 (see Fig. 3.5).

3.4.3 Cell division

An ubiquitous feature of living cells is their growth and subsequent division in daughter cells. Several models have been proposed to explain how growing cells decide to divide [154–158]. Here, we consider a model wherein the cell size grows as per the differential equation

$$d\boldsymbol{v} = \left(\alpha_1 + \frac{\alpha_1 \boldsymbol{v}}{\boldsymbol{v} + v_1}\right) dt. \tag{3.35}$$

This setup encompasses both the linear growth of cell size (if $\alpha_2 = 0$ or if $v_1 = 0$) and the exponential growth (if $\boldsymbol{v} \ll v_1$). We assume that the cell divides as per a

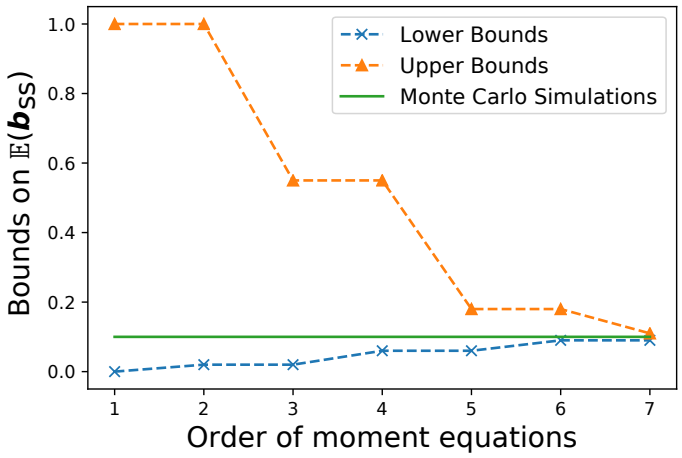

Figure 3.5: Bounds on $\mathbb{E}(\boldsymbol{\beta}_{ss})$ (i.e., the probability that the system is in the mode ss) for the TCP on-off example. The bounds improve and converge to the true value of the moment as the order of moments used in the semidefinite program is increased.

size-dependent rate

$$\lambda(\boldsymbol{v}) = (\boldsymbol{v}/v_2)^n \tag{3.36}$$

This rate is analogous to the so-called sizer strategy in the limit when $n \to \infty$ wherein the cell divides as it attains a critical volume v_2. A finite value of n represents imperfect implementation of a sizer model. Upon the reset, the cell size is reset to

$$\phi(v) = \frac{\boldsymbol{v}}{2}. \tag{3.37}$$

Since the dynamics contains a rational function, we define a new state $\boldsymbol{y} = \frac{1}{\boldsymbol{v}+v_1}$. The SHS can then be recasted as polynomial SHS with the new continuous dynamics

$$d\boldsymbol{v} = (\alpha_1 + \alpha_1 \boldsymbol{v}\boldsymbol{y})\, dt, \tag{3.38}$$

and an algebraic constraint

$$\boldsymbol{v}\boldsymbol{y} + v_2\boldsymbol{y} - 1 = 0. \tag{3.39}$$

The dynamics of the moment of a form $\mathbb{E}(\boldsymbol{v}^{m_1})$ can be computed as

$$\frac{d\mathbb{E}(\boldsymbol{v}^{m_1})}{dt} = m_1\alpha_1\mathbb{E}(\boldsymbol{v}^{m_1-1}) + m_1\alpha_2\mathbb{E}(\boldsymbol{v}^{m_1}\boldsymbol{y}) - \frac{2^m-1}{2^m v_2^n}\mathbb{E}\left(\boldsymbol{v}^{m_1+n}\boldsymbol{y}^{m_2}\right). \tag{3.40}$$

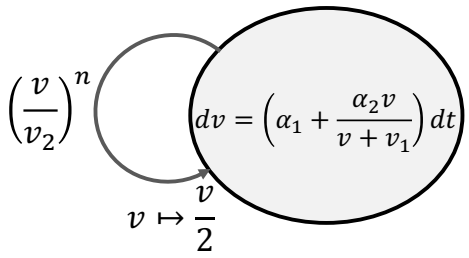

Figure 3.6: Stochastic Hybrid System representation of sizer model of cell division. The cell size v grows as per the deterministic differential equation that is a combination of two growth regimes, a linear growth with parameter α_1 and a saturating exponential growth with parameter α_2. The cell divides with intensity $(\boldsymbol{v}/v_2)^n$ and the size resets to $\boldsymbol{v}/2$ (i.e., size divides in two daughters). The parameter n represents imperfect implementation of the sizer and the cell divides at attainment of volume v_2 as $n \to \infty$.

These moment equations can be used along with the semidefinite constraints obtained from joint moments of the form $\mathbb{E}(\boldsymbol{v}^{m_1}\boldsymbol{y}^{m_2})$ and utilizing the algebraic constraints $\boldsymbol{v}\boldsymbol{y} + v_2\boldsymbol{y} - 1 = 0$.

As in the previous example, here too we can solve the steady-state moment equations. The technique can be used to explore the effect of parameters in noise in cell size. To this end, we plot the noise in cell size as a function of the cell size exponent n in Fig. 3.7. Our results show that the cell size noise decreases with increase in n, which is expected since the size control on when the division should take place becomes stronger. Similar results were obtained in [159], albeit for only exponential growth rate strategy and polynomial dynamics.

3.5 Estimating Characteristic Functions via Semidefinite Programming

So far, both in this chapter and the previous one, we have obtained bounds on moments of stochastic dynamical systems by exploiting the fact that moments satisfy semidefinite properties. It turns out that the characteristic function of a random variable is also a positive definite function [160–163]. This poses a question: can we exploit the positive definiteness of the characteristic function to estimate characteristic functions of stochastic dynamical systems? In this section, we attack this question

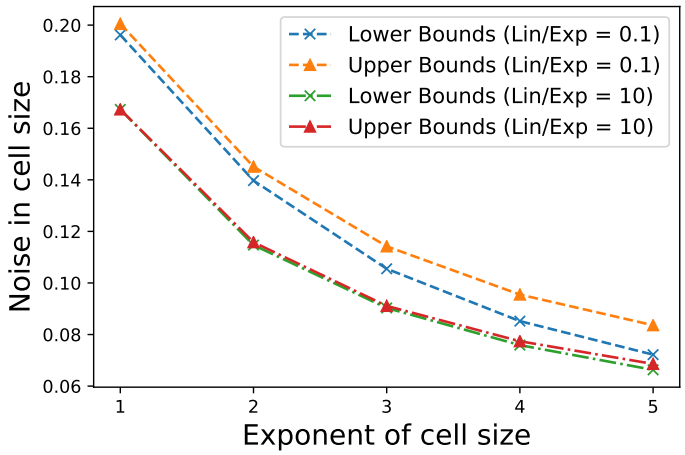

Figure 3.7: Bounds on noise (quantified via coefficient of variation squared) in cell size as a function of cell size exponent. Using ten moment equations, the bounds are computed via semidefinite program for different exponents of cell size. It is seen that the noise in cell size decreases with increase in the exponent. Moreover, the noise is lower when the linear growth coefficient α_1 is greater than the exponential growth coefficient α_2.

for a small stochastic setup consisting of one dimension. However, before going into that, we first review some basic properties of characteristic functions. In this section, we provide background results pertaining characteristic functions and their properties. The reader is referred to [160–163] for proofs and more details. For simplicity, we only consider one dimensional systems.

3.5.1 Review of Relevant Properties of Characteristic Function

The characteristic function $\varphi : \mathbb{R} \to \mathbb{C}$ for a univariate random variable $\boldsymbol{x} \in \mathbb{R}$ is defined as

$$\varphi(\omega) := \mathbb{E}\left(e^{j\omega \boldsymbol{x}}\right) = \int_{-\infty}^{\infty} e^{j\omega x} dF(x), \tag{3.41}$$

where F is the distribution function of X and $j = \sqrt{-1}$.

For any random variable, the characteristic function always exists and it uniquely determines the distribution. If $F(\boldsymbol{x})$ has a density $\rho(\boldsymbol{x})$ then $\varphi(\omega)$ can be written in

terms of ρ

$$\varphi(\omega) = \int_{-\infty}^{\infty} e^{j\omega x} \rho(\boldsymbol{x}) d\boldsymbol{x}. \tag{3.42}$$

Given $\varphi(\omega)$, the distribution function $F(\boldsymbol{x})$ and/or the density $\rho(\boldsymbol{x})$ can be obtained via inversion.

As an immediate consequence of its definition, the characteristic function of a random variable has the following properties

1. $\varphi(0) = 1$.
2. $|\varphi(\omega)| \leq 1$.
3. $\varphi(-\omega) = \overline{\varphi(\omega)}$.
4. If the random variable has a finite m^{th} order moment, then it is given by

$$\mathbb{E}\left(\boldsymbol{x}^m\right) = (-j)^m \frac{d^m \varphi}{d\omega^m}|_{\omega=0}. \tag{3.43}$$

Another important property of a characteristic function that we particularly use in this work is that it is a positive definite function. The following theorem of Bochner forms basis of our analysis.

Theorem 3 (Bochner) *A continuous complex-valued function φ on $\mathbb{R}$ is a characteristic function if and only if $\varphi(0) = 1$ and φ is positive definite [161, 162].*

The positive definiteness of a function required by this theorem is defined as follows. A complex valued function $\varphi : \mathbb{R} \to \mathbb{C}$ is said to be positive definite if the inequality

$$\sum_{k,l=1}^{q} \varphi(\omega_i - \omega_j) c_k \overline{c_l} \geq 0 \tag{3.44}$$

holds for every positive integer q, for all $\omega_1, \ldots, \omega_q \in \mathbb{R}$, and for all $c_1, \ldots, c_q \in \mathbb{C}$. In other words, φ is positive definite if and only if the matrix

$$H_c = \begin{bmatrix} \varphi(0) & \varphi(\omega_1 - \omega_2) & \ldots & \varphi(\omega_1 - \omega_q) \\ \varphi(\omega_2 - \omega_1) & \varphi(0) & \ldots & \varphi(\omega_2 - \omega_q) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(\omega_q - \omega_1) & \varphi(\omega_q - \omega_2) & \ldots & \varphi(0) \end{bmatrix} \succeq 0. \tag{3.45}$$

for an arbitrary choice of $q \in \mathbb{N}$ and $\omega_1, \ldots, \omega_q \in \mathbb{R}$.

As a consequence of this definition, a variety of properties of the characteristic function (including properties 2 and 3 above) can be established by choosing some test points and enforcing $H_c \succeq 0$. For example, consider the case $q = 1$. Then, we must have that

$$\varphi(0) \geq 0. \tag{3.46}$$

Likewise, for $q = 2$, we should have

$$\begin{bmatrix} \varphi(0) & \varphi(\omega_1 - \omega_2) \\ \varphi(\omega_2 - \omega_1) & \varphi(0) \end{bmatrix} \succeq 0. \tag{3.47}$$

Without loss of generality, we can assume $\omega_1 = 0$ and $\omega_2 = \omega \in \mathbb{R}$. Then, we should have $\varphi(-\omega) = \overline{\varphi(\omega)}$ and $|\varphi(\omega)|^2 \leq \varphi(0)^2$.

3.5.2 Characteristic Function of a Stochastic Process

Consider the SHS defined in (3.1). We restrict ourselves to a single mode SHS with no jumps, and assume that the functions f, and g are polynomials in the state $\boldsymbol{x}$. Our goal is to compute the time evolution of the test function

$$\psi(\boldsymbol{x}(t)) = e^{j\omega \boldsymbol{x}(t)}. \tag{3.48}$$

The following theorem shows how a partial differential equation governing the evolution of the characteristic function can be obtained. We wish to point out that it is presented here for a formal statement, and it has been used in some form or other in several works, e.g., see [164].

Theorem 4 *Let f and g^2 be finite polynomials of degrees $d_f \in \mathbb{N}$ and $d_g \in \mathbb{N}$ respectively. Assuming that a stationary distribution with finite moments of order $n = \max\{d_f, d_g\}$ exists, the characteristic function of the stationary distribution satisfies the following partial differential equation*

$$\frac{\partial \varphi}{\partial t} = j\omega \sum_{l=0}^{d_f} a_{f_l} j^{-l} \frac{\partial^l \varphi}{\partial \omega^l} + \frac{1}{2}(j\omega)^2 \sum_{l=0}^{d_g} a_{g_l} j^{-l} \frac{\partial^l \varphi}{\partial \omega^l}. \tag{3.49}$$

Proof: Since we assume that f, and g^2 are polynomials, without loss of generality we can take their forms to be

$$f(\boldsymbol{x}) = \sum_{l=0}^{d_f} a_{f_l} \boldsymbol{x}^l, \quad g^2(\boldsymbol{x}) = \sum_{l=0}^{d_g} a_{g_l} \boldsymbol{x}^l, \tag{3.50}$$

where $a_{f_l} \in \mathbb{R}$, and $a_{g_l} \in \mathbb{R}$ are coefficients. Taking $\psi(\boldsymbol{x}) = e^{j\omega \boldsymbol{x}}$, $\omega \in \mathbb{R}$, we have that

$$\frac{\partial^m \psi}{\partial \boldsymbol{x}^m} = (j\omega)^m \psi, \quad \frac{\partial^m \psi}{\partial \omega^m} = (j\boldsymbol{x})^m \psi \quad \forall m \in \mathbb{N} \tag{3.51}$$

Using these, we can write the extended generator as

$$\mathcal{L}\psi(\boldsymbol{x}) = j\omega\psi(\boldsymbol{x}) \sum_{l=0}^{d_f} a_{f_l} \boldsymbol{x}^l + \frac{1}{2}(j\omega)^2 \psi(\boldsymbol{x}) \sum_{l=0}^{d_g} a_{g_l} \boldsymbol{x}^l. \tag{3.52}$$

Using (3.51), the terms $\boldsymbol{x}^l \psi$ can be replaced by $(j)^{-l} \frac{\partial^l \psi}{\partial \omega^l}$. This yields

$$\mathcal{L}\psi(\boldsymbol{x}) = j\omega \sum_{l=0}^{d_f} a_{f_l} j^{-l} \frac{\partial^l \psi}{\partial \omega^l} + \frac{1}{2}(j\omega)^2 \sum_{l=0}^{d_g} a_{g_l} j^{-l} \frac{\partial^l \psi}{\partial \omega^l}. \tag{3.53}$$

Taking expectation, we get a partial differential equation in characteristic function $\varphi = \mathbb{E}(\psi(\boldsymbol{x}))$

$$\frac{\partial \varphi}{\partial t} = j\omega \sum_{l=0}^{d_f} a_{f_l} j^{-l} \frac{\partial^l \varphi}{\partial \omega^l} + \frac{1}{2}(j\omega)^2 \sum_{l=0}^{d_g} a_{g_l} j^{-l} \frac{\partial^l \varphi}{\partial \omega^l}. \tag{3.54}$$

■

If the stationary distribution with finite moments of order n exists, then we must have that $\frac{\partial \varphi}{\partial t} = 0$. This results in the ordinary differential equation (3.49). The degree of this ODE is $n = \max\{d_f, d_g\}$. It is worth pointing out that there are pathological cases wherein expectation of a test function may not be appropriately defined (see [165] for some examples), the finite moments requirement allows us to avoid them [166].

3.5.3 Solving for Stationary Characteristic Function

The ODE obtained by putting $\frac{\partial \varphi}{\partial t} = 0$ (3.49) cannot be solved analytically except for a handful of cases. However, it can be solved via numerical techniques.

Either way it would require n initial/intermediate/boundary values to find the solution. Other than the usual $\varphi(0) = 1$, previous works have either utilized prior knowledge about the system (e.g., the distribution is symmetric), or used $\lim_{|\omega|\to\infty} \varphi(\omega) = 0$ and $\lim_{|\omega|\to\infty} \frac{\partial^l \varphi(\omega)}{\partial \omega^l} = 0$ for some l [164]. In practice these are hard to incorporate in a solution. Furthermore, if one is interested only in stationary moments, then solution of $\varphi(\omega)$ only in neighborhood of zero is sufficient.

We propose a different approach to compute both the moments and the characteristic function. This approach relies on two ideas. First being the fact that the characteristic function is related with the moments as

$$\frac{\partial^l \varphi}{\partial \omega^l} \mid_{\omega=0} = j^l \mu_l, \; l = \{1, \ldots, n-1\}, \tag{3.55}$$

where $\mu_l \in \mathbb{R}$ represents the l^{th} order moment. Thus, the moments are natural quantities to be used in computing the characteristic function. Second idea is to utilize the Bochner's theorem to estimate the moments μ_l. In particular, we can use $\varphi(0) = 1$ and positive semidefinite property of the matrix M in (3.45). Using these, a semidefinite program can be formulated that gives lower and upper bounds on μ_l as stated in the theorem below.

Theorem 5 *Assuming that f and g^2 are polynomials of degree d_f and d_g, a lower bound on a moment μ_k can be obtained via the semidefinite program*

$$\min \; \mu_k \tag{3.56a}$$

$$j\omega \sum_{l=0}^{d_f} a_{f_l} j^{-l} \frac{\partial^l \varphi}{\partial \omega^l} + \frac{1}{2}(j\omega)^2 \sum_{l=0}^{d_g} a_{g_l} j^{-l} \frac{\partial^l \varphi}{\partial \omega^l} + \eta(\omega)\varphi = 0, \tag{3.56b}$$

$$\frac{\partial^l \varphi}{\partial \omega^l} \mid_{\omega=0} = j^l \mu_l, \quad l = 1, 2, \ldots n, \tag{3.56c}$$

$$\varphi(0) = 1, \tag{3.56d}$$

$$H_c \succeq 0. \tag{3.56e}$$

Here $k = \{1, \ldots, n\}$ with $n = \max\{d_f, d_g\}$ and H_c is defined as in (3.45). Further, the minimum value obtained by the program increases as size of M is increased by including more test points.

Proof: Since f and g^2 are assumed to be polynomials, Theorem 4 implies that the characteristic function $\varphi(\omega)$ satisfies the ODE of order n given by (3.56b). The moments are related with the derivatives of the characteristic function by the linear constraints in (3.56c). The constraint in (3.56d) and (3.56e) are a consequence of the Bochner's theorem. Since the objective function is linear in decision variables μ_k and the constraints are either equality or semidefinite constraints, the optimization problem is a semidefinite program [167].

Now suppose that the size of H_c is increased by including more test points $\omega_1, \ldots, \omega_q$. This corresponds to adding more constraints in the program, and the solution cannot get worse by doing so. ■

The upper bound on μ_k can be found by minimizing $-\mu_k$. Note that we can choose any test points $\omega_1, \omega_2, \ldots, \omega_q \in \mathbb{R}$ in order to generate the matrix H_c. For sake of simplicity, we will choose uniformly spaced values on the real-line. The above semidefinite program can be used to compute lower and upper bounds on each of the moments μ_k. These values can be then used to determine the solution to the ODE for characteristic function and thereby finding an approximation of the characteristic function. If the lower and upper bounds on each of the moments are reasonably close, then the approximate characteristic function would be quite close to the true characteristic function. It can be further used to compute the stationary probability density via inversion.

The formulation in (3.56) may not always be convenient, specially if (3.56b) does not have an analytical solution. In this case, we can interpret (3.56) as an optimal control problem for a linear time varying system. Specifically, consider a state vector $\chi = \begin{bmatrix} \varphi & \frac{\partial \varphi}{\partial \omega} & \cdots & \frac{\partial^{n-1}\varphi}{\partial \omega^{n-1}} \end{bmatrix}^\top$. Then, the differential equation describing the characteristic function becomes a linear time (in ω-space) varying system

$$\frac{d\chi}{d\omega} = \mathcal{P}(\omega)\chi(\omega), \tag{3.57}$$

for an appropriately defined matrix $\mathcal{P}$. In this setup, the objective would be to optimize

the elements of $\chi(0)$ subject to the linear matrix inequality (3.56e). The linear system in (3.57) can be appropriately discretized, and the decision variables of the optimization problem are values of the the state vector χ at those discrete points. The solution to this semidefinite program incurs numerical errors due to discretization.

The semidefinite program in (3.56) can be used to find bounds on first n moments where n is the degree of the ODE (3.56b). If one is interested in computing the higher order moments, the approximate characteristic function can be differentiated and computed at $\omega = 0$. By doing so, bounds on the higher order moments can also be computed. Alternatively, for systems with finite moments, one can compute the bounds on first n moments via the proposed method, and then use the fact that stationary moments are related via a linear system of equations given by

$$A\mathcal{X} + \bar{A}\overline{\mathcal{X}} = 0. \tag{3.58}$$

Here $\mathcal{X}$ is collection of moments up to some order and $\overline{\mathcal{X}}$ contains moments of order higher than those in $\mathcal{X}$. The number of elements of $\overline{\mathcal{X}}$ is as many as the degree of nonlinearity in the system given by n. While the usual moment closure methods estimate elements of $\overline{\mathcal{X}}$ in terms of those of $\mathcal{X}$, we simply supplement (3.58) with lower and upper bounds on n moments and thereby compute bounds on all other moments in (3.58).

3.5.4 Numerical Examples

Here we illustrate the proposed method for two simple examples.

3.5.4.1 Logistic Growth with Continuous Dynamics

Consider the following modified stochastic logistic growth model

$$d\boldsymbol{x} = (1 + \boldsymbol{x} - 0.1\boldsymbol{x}^2)dt + \sqrt{2}\boldsymbol{x}d\boldsymbol{w}. \tag{3.59}$$

Without the constant term in the drift, this model is widely-used in modeling population growth [168]. We added the constant term so that the trivial solution $\boldsymbol{x} = 0$ is ruled out.

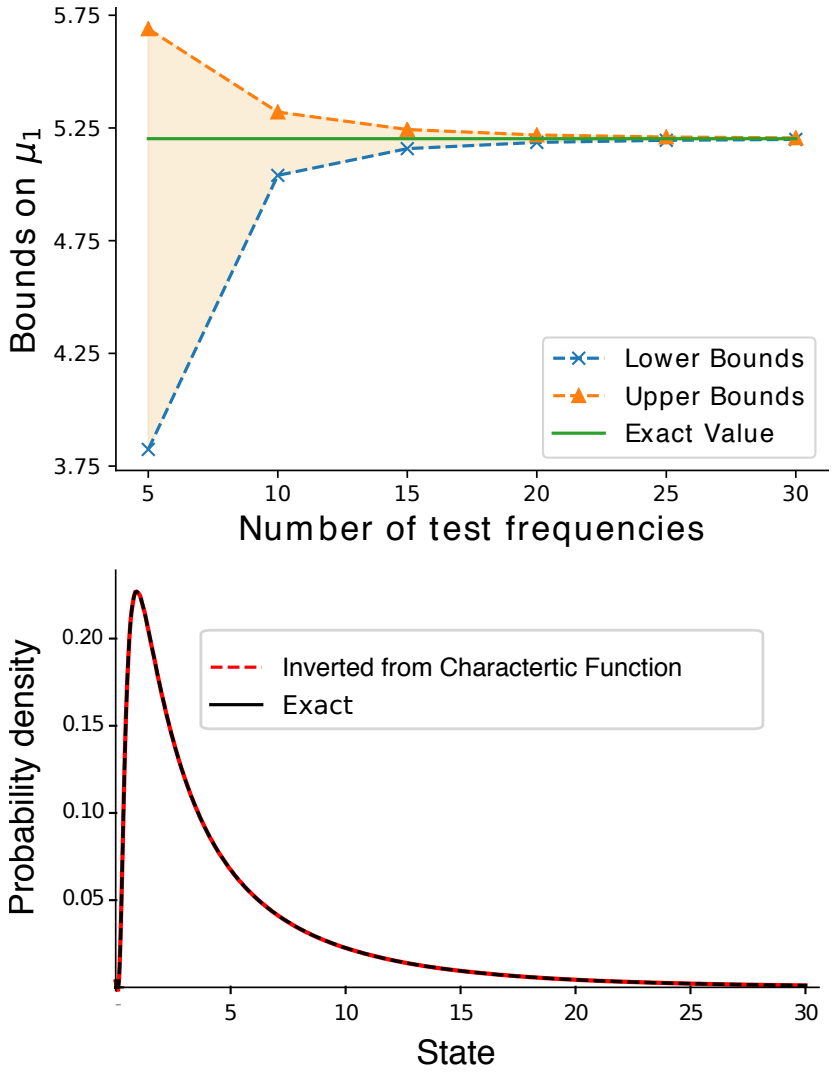

Figure 3.8: Comparison of estimated mean and characteristic function with exact values. *Top*: The mean μ_1 is estimated via semidefinite program in (3.56). The matrix H_c is constructed by taking frequencies spaced by *one* (e.g., $\omega_1 = 1, \omega_2 = 2, \ldots$). Increasing the number of test frequencies leads to a better estimation of the moment as both lower and upper bounds converge.Estimation of mean via SDP. *Bottom*: The characteristic function is estimated by using the middle value of lower and upper bounds for μ_1 obtained from 30 test frequencies in (3.62)–(3.63). The estimated probability densify function is obtained by numerically inverting the characteristic function. Notably, this particular system is exactly solvable, and the exact probability densify function is provided for comparison purpose, showing an excellent match with its estimated counterpart.

Using (3.54), the characteristic function evolves as per

$$\frac{\partial \varphi}{\partial t} = j\omega\varphi + \omega\frac{\partial \varphi}{\partial \omega} + j0.1\omega\frac{\partial^2 \varphi}{\partial \omega^2} + \omega^2\frac{\partial^2 \varphi}{\partial \omega^2}. \tag{3.60}$$

Therefore, the stationary characteristic function is the solution to the following differential equation

$$j\omega\varphi + \omega\frac{\partial \varphi}{\partial \omega} + \left(j0.1\omega + \omega^2\right)\frac{\partial^2 \varphi}{\partial \omega^2} = 0. \tag{3.61}$$

It can be shown that the above ODE has the following generalized solution

$$\varphi(\omega) = \frac{c_1}{\sqrt{5}} I_0\left(2\sqrt{0.1 - j\omega}\right) + \frac{c_2}{\sqrt{5}} K_0\left(2\sqrt{0.1 - j\omega}\right), \tag{3.62}$$

where I and K denote the modified Bessel functions of first and second kinds, and c_1, c_2 are unknown coefficients. As expected, the number of unknown coefficients is same as the order of nonlinearity in the dynamics.

To determine the coefficients, we can use the fact that $\varphi(0) = 1$ and $\frac{\partial \varphi}{\partial \omega}|_{\omega=0} = j\mu_1$, where $\mu_1 \in \mathbb{R}$ is the mean that is to be determined. This results in

$$\frac{c_1}{\sqrt{5}} I_0\left(2\sqrt{0.1}\right) + \frac{c_2}{\sqrt{5}} K_0\left(2\sqrt{0.1}\right) = 1 \tag{3.63a}$$

$$-j\sqrt{2}c_1 I_1\left(2\sqrt{0.1}\right) + j\sqrt{2}c_2 K_1\left(2\sqrt{0.1}\right) = j\mu_1 \tag{3.63b}$$

Using these, the stationary characteristic function can be written in terms of only one unknown μ_1. Now, we can compute bounds on μ_1 using the semidefinite program as in (3.56). By choosing uniformly spaced values of $\omega_1, \omega_2, \ldots, \omega_n$, we computed the maximum and minimum allowable values μ_1. We also find that increasing the size of the program by choosing more test points improves both lower and upper bounds. Taking 30 test points at $\omega_1 = 1, \ldots \omega_{30} = 30$, we get $5.2024 \leq \mu_1 \leq 5.2025$ (see Fig. 3.8, Top). This result is in excellent agreement with Monte Carlo simulations which yield an value of 5.2 for 10000 simulations.

Using the value of μ_1 obtained here, we can use the characteristic function to reconstruct the probability density function of the stationary distribution (see Fig. 3.8,

Bottom). As mentioned earlier, the bounds on μ_1 can be used to estimate bounds on higher order moments as well (results not shown here).

While the preceding discussion so far has been to consider continuous dynamics, we illustrate by a simple example below that the method also works for discrete-dynamics.

3.5.4.2 Protein Production and Decay

Consider the following reactions, where proteins are synthesized in geometric-distributed bursts of size $\boldsymbol{B}_i$ at a rate k_p, and two protein molecules come together for degradation with rate γ_p

$$\emptyset \xrightarrow{k_p} \boldsymbol{B}_i \times Protein, \quad Protein + Protein \xrightarrow{\gamma_p} \emptyset. \tag{3.64}$$

The reaction propensity functions are assumed to be k_p and $\gamma_p \boldsymbol{x}^2$, respectively, where $\boldsymbol{x}$ denotes the single-cell protein level. Taking the Fourier transform of its CME, it can be shown that the characteristic function evolves as per

$$\frac{\partial \varphi}{\partial t} = k_p \psi \left(\varphi_{\boldsymbol{B}} - 1\right) - \gamma_p (e^{-j\omega} - 1) \frac{\partial^2 \varphi}{\partial \omega^2}, \tag{3.65}$$

where $\psi_{\boldsymbol{B}} = \left\langle e^{j\omega \boldsymbol{B}} \right\rangle$ is the characteristic function of the burst size $\boldsymbol{B}$ and is given by $\varphi_{\boldsymbol{B}} = \frac{1}{1+\langle \boldsymbol{B} \rangle - \langle \boldsymbol{B} \rangle e^{j\omega}}$. Although the above partial differential equation is difficult to solve, it is possible to compute the stationary characteristic function by setting $\frac{\partial \varphi}{\partial t} = 0$ and solving the resulting ordinary differential equation

$$\varphi(\omega) = c_1 G^{2,0}_{2,2}\left(\frac{\langle U \rangle e^{j\omega}}{\langle U \rangle + 1} \,\middle|\, \begin{matrix} 1 - j\sqrt{\frac{k_p}{\gamma_p}}, 1 + j\sqrt{\frac{k_p}{\gamma_p}} \\ 0, 0 \end{matrix} \right) + c_2 \, {}_2F_1\left(-j\sqrt{\frac{k_p}{\gamma_p}}, j\sqrt{\frac{k_p}{\gamma_p}}; 1; \frac{\langle U \rangle e^{j\omega}}{\langle U \rangle + 1} \right) \tag{3.66}$$

Here coefficients c_1 and c_2 are unknown coefficients that need to be determined while $G^{2,0}_{2,2}$ and ${}_2F_1$ respectively denote Meijer G and Hypergeometric functions. To compute c_1 and c_2, one can use $\varphi(0) = 1$ and $\frac{\partial \varphi}{\partial \omega}|_{\omega=0} = j \langle \boldsymbol{x} \rangle$, where $\langle \boldsymbol{x} \rangle$ is the mean. Even though exact value of $\langle \boldsymbol{x} \rangle$ cannot be computed, it can be estimated to arbitrary precision via semidefinite programming subject to semidefinite constraints on moments

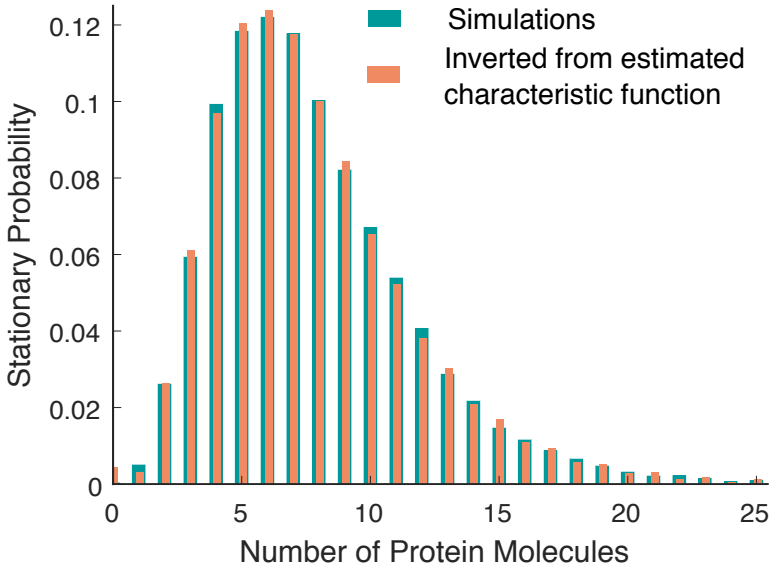

Figure 3.9: Comparison of probability mass functions computed from the proposed method and 10^5 MC simulations. The unknown moment $\langle \boldsymbol{x} \rangle$ is estimated via semidefinite with constraints in (3.45) computed at 30 frequencies equally spaced between 1 and 30, giving $7.75660 \leq \langle \boldsymbol{x} \rangle \leq 7.75667$. The characteristic function in (3.66) is computed by using $\langle \boldsymbol{x} \rangle = 7.75663$. The parameters are taken as $k_p = 10\ min^{-1}$, $\langle \boldsymbol{B} \rangle = 3$ and $\gamma_p = 0.04\ min^{-1}$.

and/or constraints encoded in $H_c \succeq 0$ for appropriately chosen (uniformly spaced, in our case) values of $\omega_1, \omega_2, \ldots, \omega_q$. Using the estimated values of c_1 and c_2 in (3.66) gives the approximate characteristic function, which in turn yields the corresponding probability mass function $\rho_k = \mathbb{P}(\boldsymbol{x} = k)$ via the inverse transform

$$\rho_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\omega k} \varphi(\omega) d\omega. \tag{3.67}$$

As shown in Fig. 3.9, comparing the estimated ρ_k with that obtained from simulations of the process shows excellent match between them.

3.6 Conclusion

In this chapter we extended the moment estimation methodology proposed in the previous chapter to a much larger class of stochastic systems called the Stochastic hybrid systems (SHSs). We first showed that the methodology straightforwardly applies to polynomial SHSs with a single-mode. We then showed that a multimode SHS can

be recasted as a single-mode SHS, thus extending the moment estimation framework to mutimode polynomial SHSs. Using a state-space augmentation, we then showed that a class of rational SHSs can also be dealt with in the same way. Finally, we proposed another semidefinite programming based method that can be used to estimate the moments as well as characteristic functions of some stochastic systems.

Chapter 4

MODELING EVENT TIMING IN SINGLE–CELLS

So far we have focused on estimating moments of species level in biochemical reaction systems. In individual cells, however, there are other quantities of physiological interest (e.g., timing of cellular events that occur upon accumulation of a regulatory protein to specific level) that are random as a consequence of the randomness in gene-product levels. For instance, lysis time of an *E. coli* cell infected by a bacteriophage λ is governed by accretion of a protein, holin, up to an effective level [169–171]. In the same vein, diploid yeast cells enter meiosis upon facing nutritional deprivation, and the decision to enter meiosis depends upon accumulation of the meiotic master regulator Ime1 [44]. Many other examples of such events appear in context of development [74–77], cell-cycle control [78–82], cell differentiation [83,84], sporulation [85,86], apoptosis [87–89], etc. As the protein levels are subject to molecular fluctuations due to inherent noise in gene expression [5–7,172–178], the timing of an event that triggers at a critical threshold is expected to exhibit cell-to-cell variation. Indeed, recent single-cell experiments have shown considerable cell-to-cell variability in timing of cellular events [169,179].

How cells are able to function reliably in face of the variability in timing? On one hand, it has been suggested that cells exploit the timing variability to their advantage by generating temporal phenotypic variability before committing to irreversible cell-fates [44]. One the other hand, precision in timing is thought to be paramount in processes related to development decisions [45]. Understanding how cells schedule key events at desired times and deal with the inherent variability in timing is critical to unravel biological design principles.

Towards this end, we formulate event timing as a first-passage time problem which corresponds to characterizing the first time at which the protein level crosses a certain threshold. Motivated from precision in timing in lysis of bacteriophage λ [169, 171], we investigate the feedback regulation strategies that can reduce noise in timing (quantified using coefficient of variation squared) around a given mean time while other gene expression parameters are fixed.

4.1 Formulating Event Timing as a First–Passage Time Problem

In this section, we provide a brief background on modeling event timing as a first-passage time problem. Let $\boldsymbol{x}(t)$ denote the protein level in a cell at a time t, where $t = 0$ represents the induction of gene expression. The event of interest is assumed to occur when the protein level crosses a specific threshold level θ (Fig. 4.1). The time to this event can be conveniently formulated as a first-passage time (FPT) problem where the timing of an event is represented by the random variable

$$\mathbf{T} := \min\{t \geq 0 : \boldsymbol{x}(t) \geq \theta | \boldsymbol{x}(0) = 0\}. \tag{4.1}$$

To compute the probability density function (pdf) of $\mathbf{T}$, we need a description of how the protein level $\boldsymbol{x}(t)$ evolves over time. For analytical tractability, we only model the protein production and degradation events. The mRNA half-life is assumed to be considerably small than that of the protein and its dynamics is approximated by a burst that follows the geometric distribution [180–182]. More precisely, the probabilities of occurrences of protein birth and death in an infinitesimal time interval $(t, t + dt)$ are considered as

$$\text{Probability}\,(x(t) = i + \boldsymbol{B} | x(t) = i) = k_i dt, \tag{4.2a}$$

$$\text{Probability}\,(x(t) = i - 1 | x(t) = i) = i\gamma dt, \tag{4.2b}$$

Here k_i is an arbitrary positive-valued function of its argument i, and it denotes the protein production rate when the protein level is $\boldsymbol{x}(t) = i$. The parameter γ denotes the degradation rate of one protein molecule. In biological terms, k_i represents a feedback regulation of the promoter activity by its own protein. Note that an increasing

(decreasing) function k_i corresponds to a positive (negative) feedback whereas a constant k_i represents absence of any feedback. Furthermore, $\boldsymbol{B}$ denotes the burst size which follows a geometric distribution

$$\mathbb{P}(\boldsymbol{B} = i) = \frac{b^i}{(b+1)^{i+1}}, \quad b \in (0, \infty), i \in \{0, 1, 2, 3 \ldots, \}, \tag{4.2c}$$

with b denoting the mean translation burst size which is equal to the translation rate divided by the mRNA degradation rate.

Note that if the protein did not decay, then $\boldsymbol{x}(t)$ accumulates over time and the $\mathbf{T}$ distribution is obtained by observing

$$\mathbb{P}\left(\boldsymbol{x}(t) \geq \theta\right) = \mathbb{P}\left(\mathbf{T} \leq t\right). \tag{4.3}$$

However, with protein degradation, the FPT calculation needs careful consideration so as to avoid counting multiple crossings of the threshold. To this end we construct an auxiliary process that has same probabilities of occurrences as $\boldsymbol{x}(t)$ in (4.2) except that the auxiliary process is absorbed as soon as the protein count crosses θ (Fig. 4.2). In this setup, the probability of the protein level reaching θ in the small time window $(t, t + dt)$ is the probability of being in state i at time t, and a jump of size $\theta - i$ or larger occurs in $(t, t + dt)$. Using the fact that for a geometrically distributed burst size $\boldsymbol{B}$,

$$\mathbb{P}\left(\boldsymbol{B} \geq \theta - i\right) = \left(\frac{b}{b+1}\right)^{\theta - i} \tag{4.4}$$

and rate of burst arrival is k_i when $x(t) = i$, the probability density function (pdf) of the first-passage time is given by

$$\rho_{\mathbf{T}}(t) = \sum_{i=0}^{\theta-1} k_i \left(\frac{b}{b+1}\right)^{\theta - i} p_i(t) dt, \tag{4.5a}$$

$$p_i(t) = \mathbb{P}\left(\boldsymbol{x}(t) = i\right) \tag{4.5b}$$

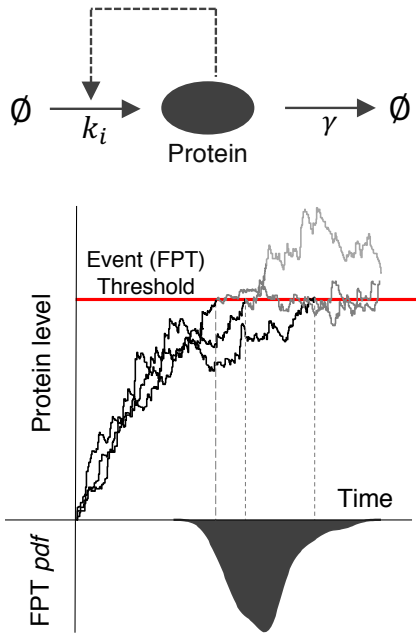

Figure 4.1: Event timing as a first-passage time problem. Top: A sketch of the gene expression model that includes production and degradation of mRNA/protein molecules. A feedback regulation scheme is implemented by assuming the protein production rate to be a function of the protein level. Bottom: An event of interest is triggered when the protein level reaches a critical event threshold for the first time. Each protein trajectory is representative of protein level over time inside individual cells. As a consequence of stochastic expression of the protein, the threshold is attained at different times in different cells. The corresponding time at which the event happens denotes the first-passage time (FPT). The portion of each trajectory after threshold has been crossed is shown in light color.

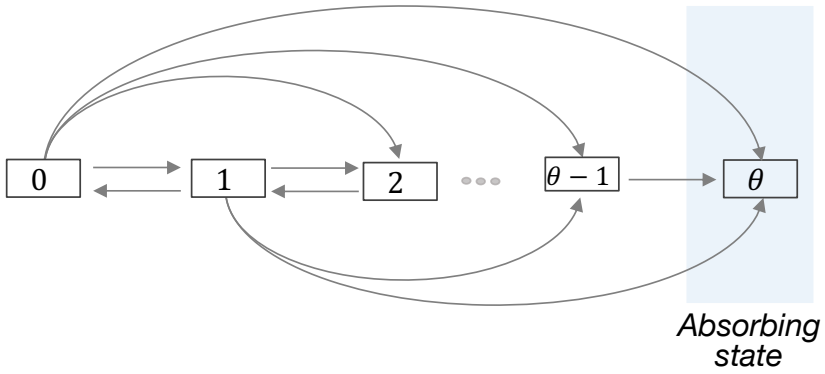

Figure 4.2: Illustration of an auxiliary process for computing the first-passage time. States $[0, 1, \ldots, \theta]$ represent the protein population counts, and arrows represent transition between states due to burst and decay events. The destination of a forward jump (a birth event) is decided by the burst size while each degradation event reduces the protein count by one. The process terminates when the protein level reaches the absorbing-state θ and the first-passage time is recorded.

The pdf can be compactly written as a product

$$\rho_{\mathbf{T}}(t) = U^{\top} P(t), \tag{4.6a}$$

$$U = \left[k_0 \left(\tfrac{b}{b+1}\right)^{\theta} \quad k_1 \left(\tfrac{b}{b+1}\right)^{\theta-1} \quad \cdots \quad k_{\theta-1} \tfrac{b}{b+1} \right]^{\top}, \tag{4.6b}$$

$$P(t) = \left[p_0(t) \quad p_1(t) \quad \cdots \quad p_{\theta-1}(t) \right]^{\top}, \tag{4.6c}$$

where U is a row vector of $k_i \mathbb{P}\left(\boldsymbol{B} \geq \theta - i\right)$ and $P(t)$ is a column vector of $p_i(t)$. The time evolution of $P(t)$ is given by the linear dynamical system

$$\frac{dP}{dt} = \Lambda P \tag{4.7}$$

derived from the Chemical Master Equations (CME) corresponding to the bursty birth-death process [102, 103]. It turns out that, in this case the matrix Λ is a Hessenberg

matrix whose i^{th} row and j^{th} column element is given by

$$a_{ij} = \begin{cases} 0, & j > i+1 \\ (i-1)\gamma, & j = i+1 \\ -k_{i-1}\frac{b}{b+1} - (i-1)\gamma, & j = i \\ k_{i-1}\frac{b^{i-j}}{(b+1)^{i-j+1}}, & j < i \end{cases} \tag{4.8}$$

$i, j \in \{1, \ldots, \theta\}$. Solving (4.7) and using (4.6a) yields the following pdf for the first-passage time

$$\rho_{\mathbf{T}}(t) = U^{\top} P(t) = U^{\top} \exp(\Lambda t) P(0), \tag{4.9}$$

where $P(0) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$ is vector of probabilities at $t = 0$ that follows from $\boldsymbol{x}(0) = 0$. While this pdf provides complete characterization of the event timing, we are particularly interested in the lower-order statistical moments of FPT. Next, we exploit the structure of the Λ matrix to obtain analytical formulas for the first and second order moments of the first-passage time.

4.2 Moments of First–Passage Time

From (4.9), the m^{th} order uncentered moment of the first-passage time is given by

$$\langle \mathbf{T}^m \rangle = U^{\top} \left(\int_0^{\infty} t^m \exp(\Lambda t) \right) P(0). \tag{4.10}$$

Here in computing the above integral, we use the fact that the matrix A is full-rank with negative eigenvalues (see Appendix B). This results in

$$\langle \mathbf{T}^m \rangle = (-1)^{m+1} m!\, U^{\top} \left(A^{-1}\right)^{m+1} P(0). \tag{4.11}$$

It turns out that the above expression of FPT moments can be written in terms of series summations. Specifically, the first two moments have the following forms

$$\langle \mathbf{T}^m \rangle = \frac{1}{k_0} + \frac{1}{b} \sum_{i=1}^{\theta} \frac{k_{i-1}}{k_{i-1} + (i-1)\gamma} \left(\frac{1}{k_{i-1}} + \sum_{j=i+1}^{\theta} \frac{1}{k_{j-1}} \prod_{l=i}^{j-1} \frac{k_l}{k_l + l\gamma} \frac{(b+1)l\gamma}{b k_{l-1}} \right), \tag{4.12a}$$

$$\left\langle \mathbf{T}^2 \right\rangle = \frac{2}{b^2} \sum_{i=1}^{\theta} \frac{k_{i-1}}{k_{i-1} + (i-1)\gamma} \left(\frac{b\,\Xi_i + \sum_{j=1}^{i} \Xi_j}{k_{i-1}} + \sum_{r=i+1}^{\theta} \frac{b\,\Xi_i + \sum_{j=1}^{r} \Xi_j}{k_{r-1}} \prod_{l=i}^{r-1} \frac{k_l}{k_l + l\gamma} \frac{(b+1)l\gamma}{bk_{l-1}} \right), \quad (4.12b)$$

where the terms Ξ_i are given by

$$\Xi_i = \frac{b}{k_0}\delta_{i-1} + \frac{k_{i-1}}{k_{i-1} + (i-1)\gamma} \left(\frac{1}{k_{i-1}} + \sum_{j=i+1}^{\theta} \frac{1}{k_{j-1}} \prod_{l=i}^{j-1} \frac{k_l}{k_l + l\gamma} \frac{(b+1)l\gamma}{bk_{l-1}} \right), \quad (4.12c)$$

with δ_{i-1} denoting the Kronecker delta which is one if $i = 1$ and zero otherwise. To our knowledge, these results represent the first analytical computations of the FPT statistics for a bursty-birth death process with a random burst size and a state-dependent burst arrival rate (i.e., feedback regulation in transcription). In what follows, we analyze the pdf and moments of $\mathbf{T}$ to explore how different model parameters affect these quantities.

4.2.1 When the Protein is Stable

Analysis of the FPT moments in some limiting cases gives important insights. For the simplest case of a stable protein ($\gamma = 0$) and a constant transcription rate (no feedback; $k_i = k$), the moment expressions simplify to

$$\langle \mathbf{T} \rangle = \frac{1}{k}\left(\frac{\theta}{b} + 1\right) \approx \frac{\theta}{bk}, \quad CV^2_{\mathbf{T}} = \frac{b^2 + X + 2b\theta}{(b+\theta)^2} \approx \frac{1+2b}{\theta}, \quad (4.13)$$

where $CV^2_{\mathbf{T}}$ represents the noise in FPT as quantified by its coefficient of variation squared (variance/mean^2; $\langle \mathbf{T}^2 \rangle / \langle \mathbf{T} \rangle^2 - 1$). The approximate formulas in (4.13) are valid for a high event threshold compared to the mean protein burst size ($\theta/b \gg 1$). The mean FPT formula can be interpreted as the time taken to reach θ with an accumulation rate bk. Further, θ/b represents the average number of burst events required for the protein level to cross the threshold, and increasing θ/b leads to noise reduction through more efficient averaging of the bursty process. One can also gain important insights, such as, the noise in FPT is invariant of the transcription rate k. Therefore, $\langle \mathbf{T} \rangle$ and $CV^2_{\mathbf{T}}$ can be independently tuned – increasing the event threshold

and/or reducing the burst size will lower the noise level. Once $CV^2_{\mathbf{T}}$ is sufficiently reduced, k can be altered to obtain a desired mean event timing.

4.2.2 When Protein is Unstable

We next investigate how changing the degradation rate influences FPT statistics.

4.2.2.1 Scale Invariance of FPT Distribution

For simplicity, we assume that there is no feedback regulation, i.e., $k_i = k$. In the deterministic sense, the protein level is described via the following ordinary differential equation

$$\frac{dx}{dt} = kb - \gamma x. \tag{4.14}$$

The solution is of the form $x = \frac{kb}{\gamma}\left(1 - e^{-\gamma t}\right)$, and it approaches the steady state $x_{ss} = kb/\gamma$ as $t \to \infty$. The timescale of protein turnover is set by the protein degradation rate γ.

In view of this, we define a normalized first-passage time $\overline{\mathbf{T}} = \gamma\,\mathbf{T}$. The pdf of $\overline{\mathbf{T}}$ can be written in terms of pdf of $\mathbf{T}$ as

$$\rho_{\overline{\mathbf{T}}}(\tau) = \frac{1}{\gamma}\,\rho_{\mathbf{T}}\left(\frac{\tau}{\gamma}\right). \tag{4.15}$$

Using the expression of $\rho_{\mathbf{T}}(t)$, we have

$$\rho_{\overline{\mathbf{T}}}(\tau) = \frac{U^\top}{\gamma}\exp\left(\frac{A\tau}{\gamma}\right)P(0). \tag{4.16}$$

Recall the forms of $U^\top$ and A from (4.6a). Using the fact that $k_i = k$ for no feedback regulation, it can be seen that elements of both A/γ and $U^\top/\gamma$ involve the ratio k/γ. As a consequence, the pdf and the moments of $\overline{\mathbf{T}}$ only depend upon the ratio k/γ and not on the individual values of k and γ. The ratio k/γ represents the average number of transcription events in occurred in a protein's lifetime and we will refer to it as the *relative transcription frequency*.

Since $\overline{\mathbf{T}} = \gamma\,\mathbf{T}$, we have that

$$\left\langle \overline{\mathbf{T}} \right\rangle = \gamma\left\langle \mathbf{T} \right\rangle, \quad \frac{\left\langle \overline{\mathbf{T}}^{\,j} \right\rangle}{\left\langle \overline{\mathbf{T}} \right\rangle^{j}} = \frac{\left\langle \mathbf{T}^j \right\rangle}{\left\langle \mathbf{T} \right\rangle^{j}}, \quad j = 2, 3, \ldots \tag{4.17}$$

Therefore, if the relative transcription frequency, k/γ, is kept constant then $\langle \mathbf{T} \rangle$ is proportional to $1/\gamma$. Other statistical metrics of $\overline{\mathbf{T}}$ and $\mathbf{T}$ that involve a higher order moment scaled by appropriate power of mean (e.g., coefficient of variation, skewness, kurtosis, etc.) are equal and they only depend on k/γ. Such behavior is referred to as scale invariance. Quite interestingly, it has recently been observed in distributions of intra– and inter–species body sizes in several species [4, 183].

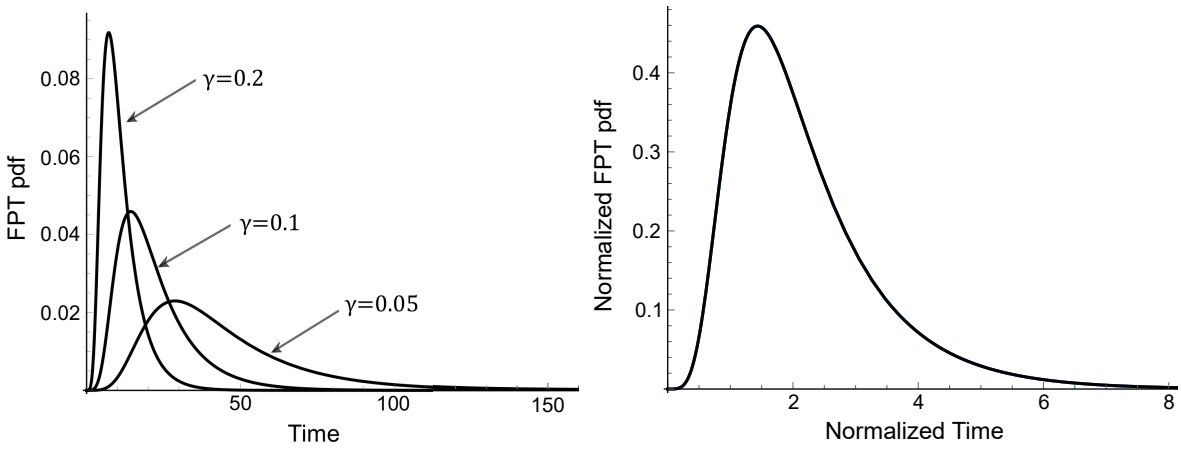

Figure 4.3: First-passage time distribution is scale invariant with respect to changes in protein degradation rate. Left: The FPT pdf is plotted for various values of protein degradation rate while keeping the ratio k/γ constant. Right: Upon normalizing the time by the degradation rate, the distributions correspond to (4.16) and collapse upon each other. The parameter values used for the plots are $k/\gamma = 30$, $b = 1$, and $\theta = 30$.

The scale invariance of FPT distribution implies that the noise (quantified using coefficient of variation squared, $CV^2_{\mathbf{T}}$) in timing could be tuned by choosing appropriate values of parameters, and the mean event time could be set by selecting the γ such that the relative transcription frequency is constant. It is worth noting that for proteins without active degradation (proteolysis), the term γ represents the dilution due to cell growth. Consequently, the mean event time is set by the cell cycle time if the other parameters are kept constant. However, for proteins with active degradation, γ could be tuned to maintain a desired k/γ. In this case, a higher value of γ requires a higher value of k and thus imposes a higher cost of protein production for the cell.

4.2.2.2 Effect of Model Parameters on FPT Statistics

A natural question that arises now is how these quantities alter with changes in the ratio k/γ? To answer this, we use the formulas in (4.12) to plot the changes in $\gamma \langle \mathbf{T} \rangle$ and $CV^2_{\mathbf{T}}$ as the steady state protein level is altered using k/γ. It is seen that for fixed values of other parameters, both $\gamma \langle \mathbf{T} \rangle$ and $CV^2_{\mathbf{T}}$ decrease with increase in k/γ (Fig. 4.4 [A]). Specifically for values of k/γ such that the steady state protein level ($x_{ss} = kb/\gamma$) is smaller than the event threshold (θ), we have $\gamma \langle \mathbf{T} \rangle \gg 1$ and $CV^2_{\mathbf{T}} \approx 1$. Thus the threshold crossing takes place at exponentially distributed random times and the time scale of the event is much larger than the time scale of protein turnover. Intuitively the exponential distribution arises because at the steady state protein level, the transient fluctuations from the steady state die out and threshold crossing becomes rare. The quantities $\gamma \langle \mathbf{T} \rangle$ and $CV^2_{\mathbf{T}}$ decrease sharply as the x_{ss} is raised towards θ; the decrease is not as sharp later on for higher values of x_{ss}.

Another point to note is that if another value of the mean burst size b is taken then the qualitative behavior of $\gamma \langle \mathbf{T} \rangle$ and $CV^2_{\mathbf{T}}$ does not change. Interestingly, comparing across different values of burst sizes reveals that when x_{ss} is smaller than θ, having a larger burst size reduces both the normalized FPT mean and noise. In contrast, when x_{ss} is larger than θ, then a higher mean burst size implies higher value of noise. This behavior could be explained by noting that when the bursting is high, it leads to higher noise among protein trajectories for a given steady-state. Because in $x_{ss} \ll \theta$ regime, the threshold crossing is noise driven, a higher noise may suppress noise in timing vis-à-vis noise in timing when the bursting is absent ($b \to 0$). In other case when x_{ss} is sufficiently higher than θ, bursting increases noise in timing and there is a one-to-one correspondence between noise in protein trajectories and noise in timing. The noise finally approaches a limiting value which corresponds to the case when k/γ is large. Note that the formulas for the FPT moments in this case reduce to

$$\gamma \langle \mathbf{T} \rangle = \frac{\gamma}{k}\left(\frac{\theta}{b}+1\right) \approx \frac{\gamma\theta}{bk}, \quad CV^2_{\mathbf{T}} = \frac{b^2+\theta+2b\theta}{(b+\theta)^2} \approx \frac{1+2b}{\theta}. \tag{4.18}$$

Thus for the no bursting case, the noise approaches the Poisson limit $1/\theta$ whereas when

burst is present it approaches $(1 + 2b)/\theta$.

Motivated from the distinct effect of the burst size, we next investigate the case when the steady-state is varied by only changing the burst size and the ratio k/γ is kept fixed. In this case, it is seen that for a given event threshold θ, there is an optimal steady state level (i.e., optimal burst size) that minimizes noise in timing (Fig. 4.4[B]). This behavior could be explained by the relative positions of event threshold and the steady state. When the event threshold is higher than x_{ss} then the noise in timing is high and it decreases as b is increased. However, towards the other limiting case when the burst size becomes high such that only a few bursting events are required to cross the threshold then the noise becomes high again. As a result, the noise is minimized by an intermediate optimal burst size. Furthermore, consistent with observations in Fig. 4.4 [A], increasing k/γ reduces both $\gamma \langle \mathbf{T} \rangle$ and $CV^2_{\mathbf{T}}$.

Until now we have considered a fixed event threshold and varied the steady-state protein level. Here, we take up the opposite scenario wherein the steady-state is constant and an appropriate event threshold is to be chosen. As shown in Fig. 4.4 [C], the noise in timing exhibits a U-shape profile as θ is varied. To understand the U shape behavior, let us consider the curve corresponding to the no bursting ($b \to 0$) case. In this case, the noise in timing approaches follows as $x_{ss} \gg \theta$. Therefore for small values of θ, the noise in timing decreases as θ increases. In contrast, as θ gets close to k/γ, crossing the threshold is dominated by noise, and therefore, becomes a rare event which results in $CV^2_{\mathbf{T}} \to 1$. These two opposite behaviors are balanced at some intermediate value of θ (close to half of k/γ) for which $CV^2_{\mathbf{T}}$ is minimum. Next, if bursts are considered, then the burst size increases noise in timing in the regime when $\theta \ll x_{ss}$, whereas it decreases noise in the regime when is θ is higher than x_{ss}. This effect can be explained by similar arguments as in Fig. 4.4 [A]–[B] that bursting increase noise in protein level and it has differential effect on noise in timing depending upon the relative positions of the event threshold and the steady state protein level.

How does the minimum noise obtained by choosing the optimal threshold varies with alterations in the steady state protein level? As shown in Fig. 4.4 [D], the optimal

noise initially decreases as the steady state is increased regardless of the underlying tuning parameter (b or k/γ). However, after a certain steady state level, the noise does not change appreciably if the tuning is carried out via the mean burst size whereas it keeps decreasing if tuning is done via k/γ. Taken together, these results show that the relative position of the steady state with respect to the event threshold is an important factor. If the steady state protein level is much higher, then the noise in protein trajectories is propagated to noise in timing. However, if the steady state level is lower in comparison to the event threshold, a higher noise in protein trajectory can even reduce noise in timing.

4.3 Optimal Feedback Strategy

Having derived the FPT moments, we investigate optimal forms of transcriptional feedback that schedule an event at a given time with the lowest $CV^2_{\mathbf{T}}$. Because $\langle \mathbf{T} \rangle$ is assumed to be fixed, minimizing $CV^2_{\mathbf{T}}$ is equivalent to minimizing $\langle \mathbf{T}^2 \rangle$. Thus, the problem mathematically corresponds to a constraint optimization problem: find transcription rates $k_0, k_1, \cdots, k_{\theta-1}$ that minimize $\langle \mathbf{T}^2 \rangle$ for a fixed $\langle \mathbf{T} \rangle$. We first consider a stable protein whose half-life is much longer than the event timescale, and hence, degradation can be ignored ($\gamma = 0$).

4.3.1 Optimal Feedback for a Stable Protein

When the protein of interest does not decay ($\gamma = 0$), the expressions for the FPT moments take much simpler forms

$$\langle \mathbf{T} \rangle = \frac{1}{k_0} + \frac{1}{b} \sum_{i=0}^{\theta-1} \frac{1}{k_i}, \tag{4.19a}$$

$$\langle \mathbf{T}^2 \rangle = \frac{2}{b^2} \left(\frac{\tau_0}{b k_0} + \sum_{i=0}^{\theta-1} \frac{\tau_i}{k_i} \right), \ \tau_i := \frac{b}{k_i} + \sum_{j=i}^{\theta-1} \frac{1}{k_j}. \tag{4.19b}$$

Note that in (4.19a) the contribution of k_0 (transcription rate when there is no protein) differs from the other transcription rates k_i, $i \in \{1, 2, \cdots, \theta - 1\}$. For instance, when the event threshold is large compared to the mean burst size ($\theta \gg b$), then the term

$1/k_0$ can be ignored and $\langle \mathbf{T} \rangle \approx \sum_{i=0}^{\theta-1} 1/bk_i$. In contrast, if the burst size is large ($b \gg \theta$) then $\langle \mathbf{T} \rangle \approx 1/k_0$, as a single burst event starting from zero protein molecules is sufficient for threshold crossing. Similar observation for different contributions of k_0 can be made about (4.19b).

It turns out that, for these simplified formulas, the problem of minimizing $\langle \mathbf{T}^2 \rangle$ given $\langle \mathbf{T} \rangle$ can be solved analytically using the method of Lagrange multipliers (see Appendix B). The optimal transcription rates are given by

$$k_0 = \frac{1+b}{1+2b} \frac{2b+\theta}{b \langle \mathbf{T} \rangle}, \quad k_i = \frac{1+2b}{1+b} k_0, \quad 1 \leq i \leq \theta - 1, \tag{4.20}$$

and all rates are equal to each other except for k_0. Intuitively, the difference for k_0 comes from the fact that it contributes differently to the FPT moments as compared to other rates. Note that for a small mean burst size ($b \ll 1$), $k_0 = k_i$, whereas $k_0 = k_i/2$ for a sufficiently large b. Despite this slight deviation in k_0, for the purposes of practical implementation, the optimal feedback strategy in this case is to have a constant transcription rate (i.e., no feedback in protein expression).

We tested the above result for a more complex stochastic gene expression model that explicitly includes mRNA dynamics via Monte Carlo simulations (Fig. 4.5). For ease of implementation, the feedbacks are assumed to be linear

$$k_i = c_1 + c_2 i, \quad i \in \{0, 1, \ldots\} \tag{4.21}$$

where $c_2 = 0$ represents a no feedback, and $c_2 > 0\,(c_2 < 0)$ denotes a positive (negative) feedback. In agreement with (4.20), a no-feedback strategy outperforms negative/positive feedbacks in terms of minimizing noise in FPT around a given mean event time. The qualitative shape of trajectories in Fig. 4.5 is determined by the feedback strategy employed, with no feedback resulting in linear time evolution of protein levels. This provides an intriguing geometric interpretation of our results – an approximate linear path from zero protein molecules at $t = 0$ to θ molecules at time $\langle \mathbf{T} \rangle$ provides the highest precision in timing. Next, we discuss the optimal feedback strategy when protein degradation is taken into consideration.

4.3.2 Optimal Feedback for an Unstable Protein

Now consider the scenario where protein degradation cannot be ignored over the event timescale ($\gamma \neq 0$). Unfortunately, the expressions of the FPT moments are too convoluted for the optimization problem to be solved analytically, and the effect of different feedbacks is needed to be investigated numerically.

We implement the feedbacks using physiologically-relevant Hill functions, where the transcription rates for a negative feedback mechanism take the following form

$$k_i = \frac{k_{\max}}{1 + (ci)^H}, \quad i \in \{0, 1, \ldots\}. \tag{4.22}$$

Here H denotes the Hill coefficient, $k_{\max}$ corresponds to the maximum transcription rate, and c characterizes the negative feedback strength with $c = 0$ representing no feedback [184, 185]. Similarly, a positive feedback is assumed to take the form

$$k_i = k_{\max}\left(r + (1-r)\frac{(ci)^H}{1 + (ci)^H}\right) = k_{\max}\frac{r + (ci)^H}{1 + (ci)^H}. \tag{4.23}$$

Note that an additional parameter $r \in (0, 1)$, referred to as the *basal strength* is introduced in (4.23). This is to ensure that the transcription rate in protein absence $k_0 = k_{\max} r > 0$, and this is necessary to prevent protein levels from getting stuck at zero molecules.

To find the optimal feedback mechanism, our strategy is as follows: for given r and H, choose a certain feedback strength c in (4.22)/(4.23), appropriately tune $k_{\max}$ for the desired mean event timing, and explore the corresponding noise in FPT as measured by its coefficient of variation squared $CV^2_{\mathbf{T}}$. Counter-intuitively, results show that for a given value of γ, a negative feedback loop in gene expression has the highest $CV^2_{\mathbf{T}}$, and its performance deteriorates with increasing feedback strength (Fig. 4.6 (Top)). In contrast, $CV^2_{\mathbf{T}}$ first decreases with increasing strength of the positive feedback, and then increases after an optimal feedback strength is crossed (Fig. 4.6 (Top)). Thus, when the protein is not stable, precision in timing is attained by having a positive feedback in protein synthesis with an intermediate strength.

We next explore how the minimal achievable noise in event timing, for a fixed $\langle \mathbf{T} \rangle$, varies with the protein decay rate γ. Our analysis shows that for a given basal

strength r, the minimum $CV_{\mathbf{T}}^2$ obtained via positive feedback increases monotonically with γ, and $CV_{\mathbf{T}}^2 \to 1$ as γ becomes large (Fig. 4.6 (Bottom)). A couple of interesting observations can be made from Fig. 4.6 (Bottom): i) The difference in $CV_{\mathbf{T}}^2$ for optimal feedback and no feedback is indistinguishable when the protein is stable ($\gamma = 0$) or highly unstable ($\gamma \to \infty$); ii) For a range of intermediate protein half-lives, the optimal feedback strategy provides better reduction of CV_T^2, as compared to no feedback regulation (which also corresponds to minimum $CV_{\mathbf{T}}^2$ obtained via a negative feedback); iii) Lowering the basal strength r results in better performance in terms of noise suppression; and iv) A linear feedback based on (4.21) outperforms feedbacks based on Hill functions, and provides significantly lower levels of $CV_{\mathbf{T}}^2$ for high protein decay rates.

Why is positive feedback the optimal control strategy for ensuring precision in event timing? One way to understand this result is to consider the linear feedback form (4.21), in which case the mean protein levels evolve according to the following ordinary differential equation

$$\frac{dx(t)}{dt} = b(c_1 + c_2 x) - \gamma x, \quad x(0) = 0. \tag{4.24}$$

Recall the geometric argument presented in Fig. 4.5, where an approximately linear path for the protein to reach the prescribed threshold in a given time provides the highest precision in event timing. While no feedback ($c_2 = 0$), and negative feedback ($c_2 < 0$) in (4.24) will create nonlinear protein trajectories, choosing a positive value $c_2 \approx \gamma/b$ results in linear $x(t)$, and hence minimal noise in event timing. Indeed, our detailed stochastic analysis shows that the optimal feedback strength that minimizes $CV_{\mathbf{T}}^2$ in the stochastic model is qualitatively similar to $c_2 \approx \gamma/b$.

4.4 Biological Implications and Discussion

Biological timers are ubiquitous at different levels of abstraction in biology [186]. At single-cell level, timers are used to determine when to respond to an internal or external cue, or commit to a certain cell fate, etc. As the timers are usually implemented via expression of a regulatory protein up to a certain threshold, the inherent stochasticity of gene expression leads to noise in the measured time. We have systematically

investigated ingredients essential for precision in timing of biochemical events at the level of single cells. Our approach relies on modeling event timing as the first-passage time for a stochastically expressed protein to cross a threshold level. This framework was used to uncover optimal strategies for synthesizing the protein that ensures a given mean time to event triggering (threshold crossing), with minimal fluctuations around the mean. The key insights can be summarized as: (i) The characteristics of event timing is mostly dictated by how the event threshold and the steady-state protein level are placed with respect to each other. (ii) If the protein half-life is much longer than the timescale of the event, the highest precision in event timing is attained by having no feedback, i.e., expressing the protein at a constant rate (Fig. 4.5); (iii) If the protein half-life is comparable or shorter than the timescale of the event, then positive feedback provides the lowest noise in event timing. Moreover, the minimum achievable noise in timing increases with the protein decay rate γ and approaches $CV_{\mathbf{T}}^2 = 1$ as $\gamma \to \infty$ (Fig. 4.6).

How robust are these findings to alternative noise sources and key modeling assumptions? For example, the model only considers noise from low-copy number fluctuation in gene product levels, and ignores any form of "extrinsic noise" that arises from cell-to-cell differences in gene expression machinery [187, 188]. To incorporate such extrinsic noise, we alter the transcription rate to $k_i\boldsymbol{\zeta}$, where $\boldsymbol{\zeta}$ is drawn from an a priori probability distribution at the start of gene expression ($t = 0$), and remains fixed till the threshold is reached. Interestingly, the optimal feedback derived in (4.20) does not change even after adding extrinsic noise to the transcription rate (see appendix B).

Another important model aspect is geometrically distributed protein burst size, which results from the assumption of exponentially distributed mRNA lifetimes. We have also explored the scenario of perfect memory in the mRNA degradation process, which results in a mRNA lifetime distribution given by the delta function. In this case, the protein burst size is Poisson and the optimal feedback strategy is fairly close to having no feedback for a stable protein (Fig. 4.7).

Next, we discuss the biological implications of our findings in the context of

phage λ's lysis times, i.e., the time taken by the virus to destroy infected bacterial cells.

4.4.1 Connecting Theoretical Insights to λ Lysis Times

Phage λ has recently emerged as a simple model system for studying event timing at the level of single cells [45, 169]. After infecting *E. coli*, λ expresses a protein, holin, which accumulates in the inner membrane. When holin reaches a critical threshold concentration, it undergoes a structural transformation, forming holes in the membrane [170]. Subsequently the cell lysis and phage progeny are released into the surrounding medium. Since hole formation and cell rupture are nearly simultaneous, lysis timing depends on de novo expression and accumulation of holin in the cell membrane up to a critical threshold [170]. Data reveals precision in the timing of lysis – individual cells infected by a single virus lyse on average at 65 mins, with a standard deviation of 3.5 mins, implying a coefficient of variation of $\approx 5\%$. Such precision is expected given the existence of an optimal lysis time [189–192]. Intuitively, if λ lysis is early then there are no viral progeny. In contrast, if λ lysis is late then the infected cell could die before lysis is effected, trapping the virus with it.

The threshold for lysis is reported to be a few thousand holin molecules [193]. Moreover, the holin mean burst size (average number of holins produced in a single mRNA lifetime) is estimated as $b \approx 1-3$ [193]. Based on our FPT moment calculations in (4.13), such a small protein burst size relative to the event threshold will yield a tight distribution of lysis times. Interestingly, (4.13) provides insights for engineering mutant λ that lyse, on average, at the same time as the wild type, but with much higher noise. This could be done by lowering the threshold θ for lysis through mutations in the holin amino acid sequence [169], and also reducing the holin mRNA transcription rate k or mean protein burst size b by reducing the translation rate so as to keep the same mean lysis time. Notably, the holin proteins are long-lived and do not degrade over relevant timescales [194], therefore λ's lysis system with no known feedback in

holin expression provides better suppression of lysis-time fluctuation compared to any feedback regulated system.

4.4.2 Additional Mechanism for Noise Buffering

The surprising ineffectiveness of feedback control motivates the need for other mechanisms to buffer noise in event timing. Intriguingly, λ uses feedforward control to regulate the timing of lysis which is implemented through two proteins with opposing functions: holin and antiholin [195, 196]. In the wild-type virus both proteins are expressed in a 2:1 ratio (for every two holins there is one antiholin) from the same mRNA through a dual start motif. Antiholin binds to holin and prevents holin from participating in hole formation, creating an incoherent feedforward circuit. Synthesis of antiholin leads to a lower burst size for active holin molecules, and increases the threshold for the total number of holins needed for lysis – both factors functioning to lower the noise in event timing. Consistent with this prediction, variants of λ lacking antiholin are experimentally observed to exhibit much higher intercellular variation in lysis times as compared to the wild-type virus [169, 171]. Succinctly put, λ encodes several regulatory mechanisms (low holin burst size; no feedback regulation; feedforward control) to ensure that single infected cells lyse at an optimal time, despite the stochastic expression of lysis proteins.

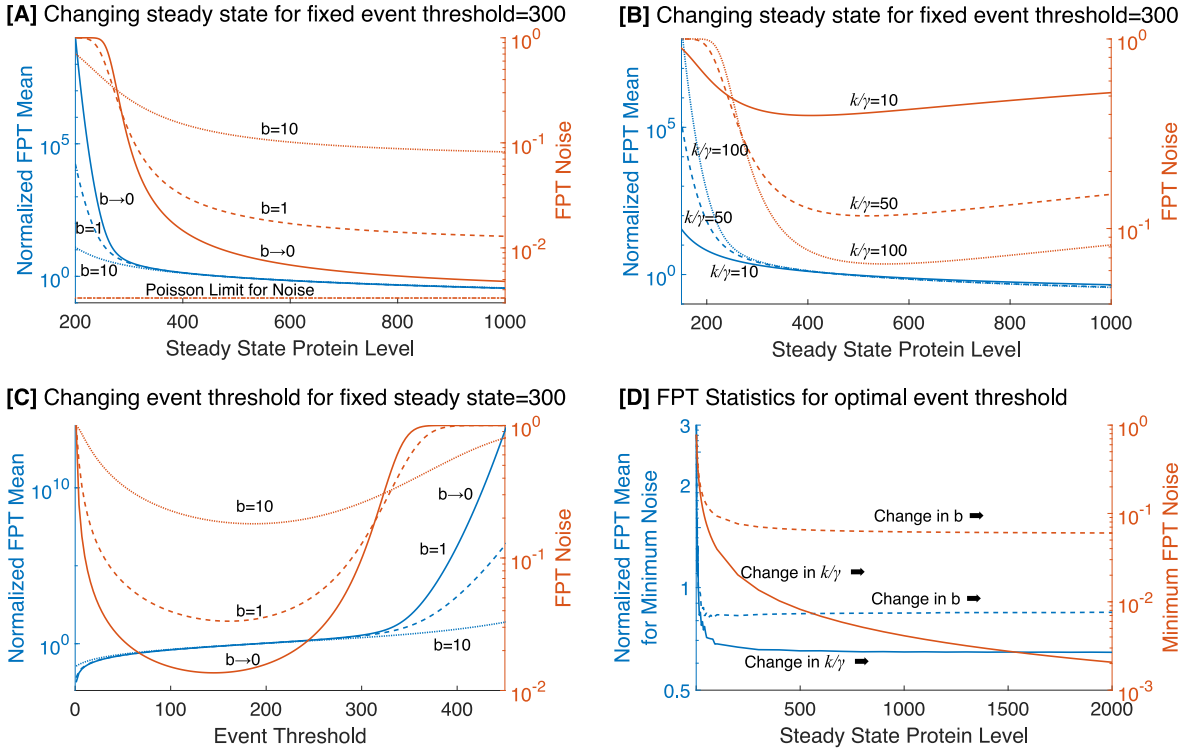

Figure 4.4: Relative positions of event threshold and steady state protein level determine FPT statistics. [A] . The Normalized mean FPT and the noise in FPT are plotted as the steady state protein level is increased by keeping the translation mean burst size, b, fixed and varying the relative transcription frequency, k/γ. [B] The Normalized mean FPT and the noise in FPT are plotted as the steady state protein level is increased by keeping k/γ fixed and varying b. [C] . Increasing the event threshold for a fixed steady-state protein level results in an increase in the normalized FPT mean and a U-shape curve for the FPT noise. [D] . The minimum noise obtained by selecting the optimal event threshold initially decreases by increase in both k/γ and b, but is less sensitive to changes in b for high steady-state level.

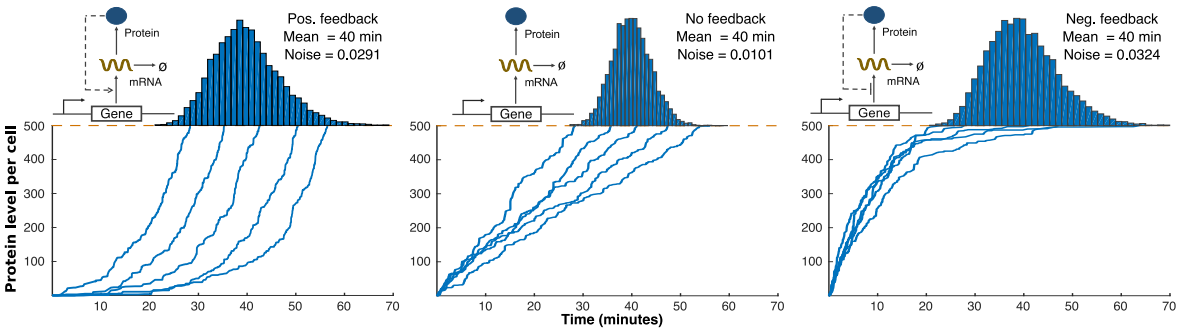

Figure 4.5: For a stable protein, no feedback provides the lowest noise in event timing for a fixed mean FPT. Protein trajectories obtained using the Stochastic Simulation Algorithm (SSA) for a stochastic gene expression model with positive feedback (left), no feedback (middle), negative feedback (right) [3]. The threshold for event timing is assumed to be 500 protein molecules and feedback is implement by assuming linear form of transcription rates as $k_i = c_1 + c_2 i$, The value of c_2 is taken as $0.05\ \text{min}^{-1}$ for positive feedback and $-0.05\ \text{min}^{-1}$ for negative feedback. For each feedback, the parameter c_1 is taken such that the mean FPT is kept constant (40 minutes). The mRNA half-life is assumed to be 2.7 min, and proteins are translated from mRNAs at a rate $0.5\,\text{min}^{-1}$, which corresponds to a mean burst size of $b = 2$. Histograms on the top represent distribution of FPT from $10,000$ Monte Carlo simulations.

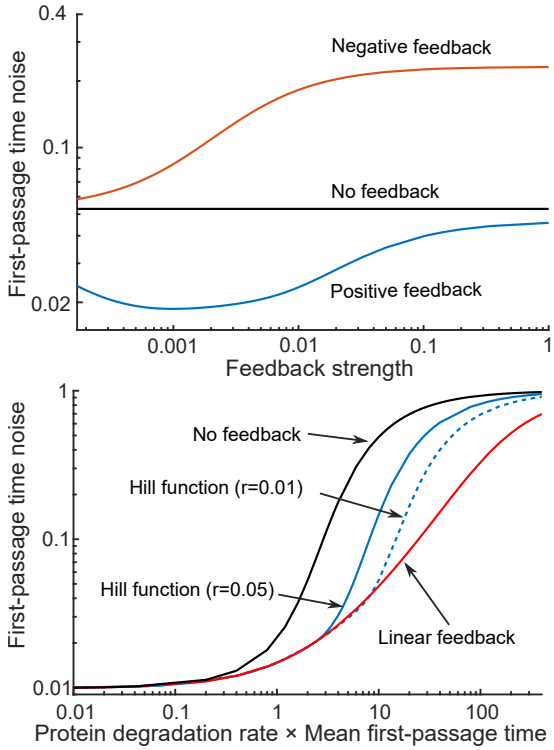

Figure 4.6: For an unstable protein, positive feedback provides the lowest noise in event timing for a fixed mean FPT. *Top*: Noise in timing ($CV^2_{\mathbf{T}}$) as a function of the feedback strength c for different control strategies. The value of $k_{\max}$ is changed in (4.22)/ (4.23) so as to keep $\langle \mathbf{T} \rangle = 40$ mins fixed. The performance of the negative feedback worsens with increasing feedback strength. In contrast, positive feedback with an optimal value of c provides the highest precision in event timing. Other parameters used are $\gamma = 0.05\,\mathrm{min}^{-1}$, $\theta = 500$ molecules, $H = 1$, $b = 2$, and for positive feedback $r = 0.05$. *Bottom*: The minimum value of $CV^2_{\mathbf{T}}$ obtained via positive feedback increases monotonically with the protein degradation rate. A smaller basal promoter strength $r = 0.01$ in (4.23) gives better noise suppression than a larger value $r = 0.05$. For comparison purposes, $CV^2_{\mathbf{T}}$ obtained without any feedback ($c = 0$), and a linear feedback with c_1 and c_2 in (4.21) chosen so at minimize $CV^2_{\mathbf{T}}$ for a given $\langle \mathbf{T} \rangle = 40$ mins are also shown. The parameters values used are $\theta = 500$ molecules, $H = 1$, and $b = 2$.

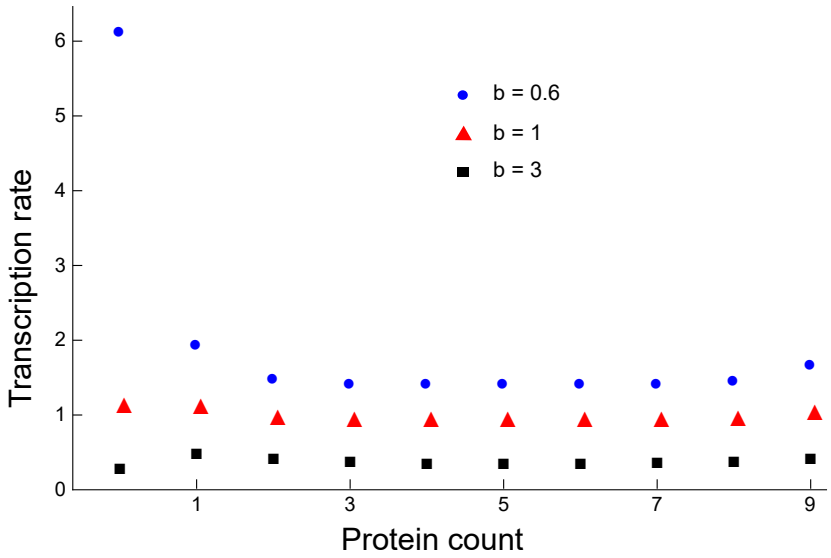

Figure 4.7: Optimal feedback strategy for Poisson distributed burst size. The optimal transcription rates obtained via numerical optimization for different values of mean burst size are shown. The event threshold is assumed to be 10 molecules, and the mean FPT is constrained to be 10 minutes.

Chapter 5

A MECHANISTIC FRAMEWORK FOR BACTERIAL CELL DIVISION TIMING

One common theme underlying life across all organisms is recurring cycles of growth of a cell, and its subsequent division into two viable progenies. How an isogenic population of proliferating cells maintains a narrow distribution of cell size, a property known as cell size homeostasis, has been extensively studied, e.g., see [4, 154–157] and references therein. From a phenomenological standpoint, recent single-cell experiments have shown that several microorganisms achieve homeostasis via what is called an adder principle: on average, a cell adds a constant size between birth and division regardless of its size at birth [4, 197–200]. Interestingly, the size accumulated between birth to division also has remarkable statistical properties: (1) not only the mean but its distribution itself is independent of cell size at birth for a given growth condition; (2) its distributions in different growth conditions collapse upon each other if scaled by their respective mean values [4]. An unresolved central issue of interest is to explore the biophysical mechanisms underpinning the adder form of cell size control.

One biophysical mechanism proposed in this regard assumes a protein acting as time-keeper between one division to next [199, 201–203]. This protein's production is started right after birth of a cell, and its production rate is proportional to instantaneous volume (size) of the cell which grows exponentially until division. The cell divides when the protein accrues up to a prescribed level after which the protein is assumed to degrade instantaneously (Fig. 5.1(a)). Along the same lines, it has also been proposed that the protein keeps time between subsequent occurrences of some event in the cell cycle other than division [82, 204–206]. While it has been shown that this simple mechanism can lead to the adder principle of cell size control in the mean

sense [82, 203], it remains to be seen whether the stochastic properties can also be explained by it.

A plausible source of stochasticity is the noise in expression of the time-keeper protein wherein randomness in transcription and translation, coupled with low copy number of species, results in significant cell-to-cell variation in protein levels [6, 7, 207–209]. As a consequence, the protein count attains a given threshold at different times in different cells, leading to variability in cell division timing (Fig. 5.1(b)). We apply the first-passage time framework to model cell division time, and consequently quantify the volume added between birth to division (Fig. 5.1(b),(c),(d)). Consistent with data from [4], the model predicts that whereas the mean cell-division time decreases as cell size at birth increases, the noise in cell-division time increases with increment in newborn cell size. Our results further show that distribution of the added volume is independent of the newborn cell size [4]. We also show that distributions of the added volume, and cell division time have scale-invariant forms: they collapse upon rescaling with their respective means in different growth conditions [4, 210]. Lastly, we discuss implications of these findings in identification of the time-keeper protein and also deliberate upon various modifications to the proposed model that result in deviations from the adder principle.

5.1 Model description

Considering a cell with volume at birth V_0, its volume at a time t after birth is given by

$$V(t) = V_0 \exp(\alpha t), \tag{5.1}$$

where α represents the cell growth rate. The time-keeper protein is assumed to be produced at a time dependent rate $r(t) = k_m V(t)$, where k_m is a proportionality constant representing the transcription rate of protein in concentration sense. We consider gene expression in the translation burst limit: each mRNA molecule degrades

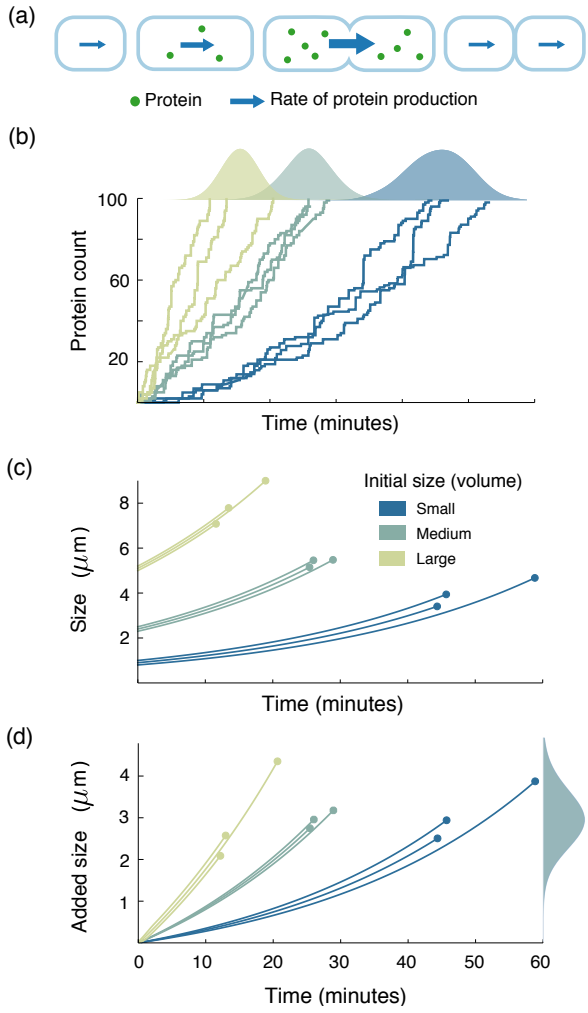

Figure 5.1: A molecular mechanism that explains adder principle. *(a)* A rod-shaped growing cell starts synthesizing a protein right after its birth. The production rate of the protein scales with the size (volume) of the cell. When the protein's copy number attains a certain level, the cell divides and the protein is degraded. *(b)* The stochastic evolution of the protein is shown for cells of three different sizes at birth. Each cell divides when the protein's level achieves a specific threshold. The distribution of FPT generated via 1000 realizations of the process for each newborn cell volume is shown above the three corresponding trajectories. The FPT distribution depends upon the newborn cell size: on average the protein in a smaller cell takes more time to reach the threshold as compared to the protein in a larger cell. *(c)* The time evolution of size is shown for cells of different initial volume. The size is assumed to grow exponentially until the protein in the top figure reaches a critical threshold. The division is assumed to take place then. *(d)* The size added to the initial size when the division event takes place is shown for cells of different volume. The distribution of the added size is independent of the initial size of the cell, thus reproducing an adder model.

instantaneously after producing a burst of protein molecules [180–182, 211–213]. Thus, the protein synthesis is given by the following biochemical reaction:

$$Gene \xrightarrow{r(t)} Gene + \boldsymbol{B}_i \times Protein. \tag{5.2}$$

In this setup, the burst arrival rate (transcription rate) scaling with the cell volume is an essential component of maintaining concentration homeostasis. Indeed such a dependency of transcription rate on cell volume has been observed in mammalian cells [214]. The variable B_i denotes the size of i^{th} protein burst which, for each $i \in \{1, 2, 3, \cdots\}$, is independently drawn from an arbitrary positive-valued distribution. It essentially represents the number of protein molecules synthesized in a single mRNA lifetime and typically follows a geometric distribution [178, 211–213, 215, 216].

Note that the protein count $\boldsymbol{x}(t)$ is a sum of independent and identically distributed protein bursts

$$\boldsymbol{x}(t) = \sum_{i=i}^{n} \boldsymbol{B}_i, \quad \langle \boldsymbol{B}_i \rangle := b, \tag{5.3}$$

where n is the number of bursts arrived (transcription events) in time interval $[0, t]$. This protein acts as a time-keeper between two division events: it starts to get produced right after birth, the cell divides when the protein level $\boldsymbol{x}(t)$ crosses a threshold θ, and the protein degrades thereafter. We would like to mention that though the model described here is for gene expression, it can be used to model any process involving accumulation of molecules wherein the production rate scales with the volume. A general distribution for $\boldsymbol{B}_i$ allows a wider range of processes to be covered. The scope can be further widened by considering the parameter k_m to be a function of the growth rate α [217, 218].

5.2 Cell Division Time as a First–Passage Time Problem

We model cell division as the time at which the protein level $\boldsymbol{x}(t)$ reaches a certain threshold θ. We use a first-passage time (FPT) framework to model cell division which can be mathematically put as

$$\mathbf{T} := \inf \{t : \boldsymbol{x}(t) \geq \theta | \boldsymbol{x}(0) = 0\}. \tag{5.4}$$

Note that as expected, $\mathbf{T}$ depends upon the cell volume at birth V_0 since $\boldsymbol{x}(t)$ is produced at a volume dependent rate. Therefore $\mathbf{T}$ above is a conditional random variable assuming a given V_0. However, we have deliberately left out any reference to it for notation convenience.

To determine the distribution of $\mathbf{T}$, we first find the distribution of the minimum number of burst (transcription) events $\mathbf{N}$ required for $\boldsymbol{x}(t)$ to reach the threshold θ. Since $\boldsymbol{x}(t)$ can only increase, distribution of $\mathbf{N}$ is related with that of burst size as

$$\mathbb{P}\left(\mathbf{N} \leq n\right) = \mathbb{P}\left(\sum_{i=1}^{n} \boldsymbol{B}_i \geq \theta\right). \tag{5.5}$$

As a specific yet physiologically relevant example, when the burst size distribution is considered to be geometric, the probability mass function of $\mathbf{N}$ denoted as $\rho_{\mathbf{N}}(n)$ is given by

$$\rho_{\mathbf{N}}(n) := \mathbb{P}\left(\mathbf{N} = n\right) = \binom{n+\theta-2}{n-1} \left(\frac{1}{b+1}\right)^{n-1} \left(\frac{b}{b+1}\right)^{\theta}, \tag{5.6}$$

where b represents the mean (or expected value) of burst size $\boldsymbol{B}_i$ [171].

Having determined $\mathbf{N}$, $\mathbf{T}$ is same as the time at which the $\mathbf{N}^{th}$ burst event occurs. Let $\mathbf{T}_n, n = 1, 2, 3, \cdots$ represent the time at which n^{th} burst event takes place, then the probability density function of $\mathbf{T}$, denoted by $\rho_{\mathbf{T}}(t)$, can be obtained by conditioning arguments:

$$\mathbb{P}\left(\mathbf{T} \in (t, t+dt) | \mathbf{N} = n\right) := \rho_{\mathbf{T}}(t|\mathbf{N} = n)dt = \rho_{\mathbf{T}_n}(t)dt \tag{5.7}$$

$$\implies \rho_{\mathbf{T}}(t) = \sum_{n=1}^{\infty} \rho_{\mathbf{T}_n}(t)\, \rho_{\mathbf{N}}(n). \tag{5.8}$$

Here $\rho_{T_n}(t)$ represents the probability density function of T_n which is governed by the underlying burst arrival process. In our case, the transcription (burst arrival) rate is time dependent which implies that the burst arrival process is an inhomogeneous Poisson process. The probability density function of the arrival times is given by the following expression [219, 220]

$$\rho_{\mathbf{T}_n}(t) = \frac{(R(t))^{n-1}}{(m-1)!} r(t) \exp(-R(t)). \tag{5.9}$$

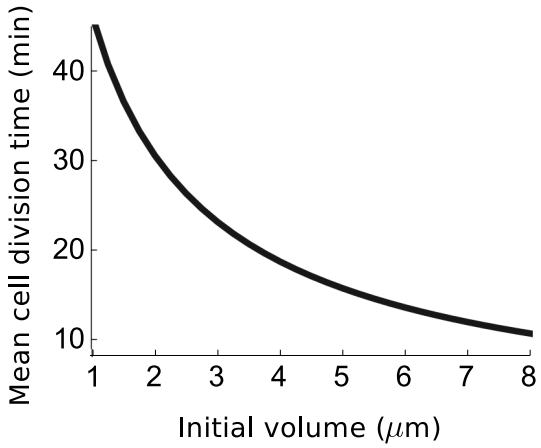

Figure 5.2: The division time decreases as the cell size at birth increases. The division time (mean FPT) given a newborn volume is computed numerically for each value of initial cell volume V_0. The protein production is assumed in geometric bursts and the parameters assumed are: $k_m = 1$ per minute, $\theta = 100$ molecules, $b = 3$ molecules, $\alpha = 0.03$ per minute.

The transcription or burst arrival rate $r(t)$ is referred to as the intensity function of the corresponding inhomogeneous Poisson process. Also, $R(t)$ is the mean value function of the inhomogeneous poisson process

$$R(t) := \int_0^t r(s)ds = \frac{k_m V_0}{\alpha}\left(e^{\alpha t} - 1\right). \tag{5.10}$$

Using (5.9) in (5.8), the expression for the probability density function of FPT for a cell of given volume V_0 at birth becomes

$$\rho_{\mathbf{T}}(t) = \sum_{n=1}^{\infty} \frac{(R(t))^{n-1}}{(n-1)!} r(t) \exp(-R(t)) \rho_N(n). \tag{5.11}$$

The FPT distribution in (5.11) qualitatively emulates the experimental observations by [4] that the mean cell division time decreases as the cell size at birth is increased (Fig. 5.2). This is intuitively expected as a cell with large size at birth will have a higher transcription rate as compared to a cell with small size. Hence, on average, the time taken by the protein to reach the prescribed threshold is smaller in the larger cell. The model also predicts that the noise (quantified using coefficient of variation squared, CV^2) in the cell division time increases as new born cell volume is increased

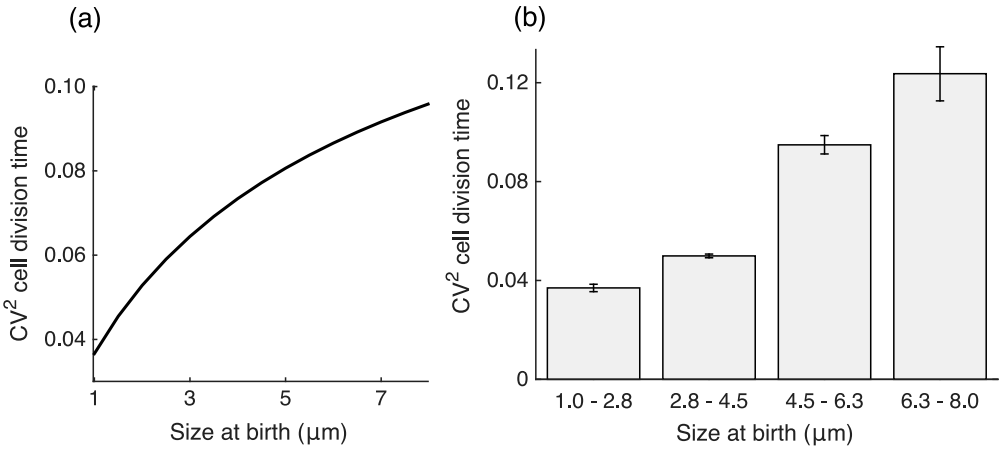

Figure 5.3: The noise in the cell division time increases with initial volume. *Left*: The noise in division time (CV^2 of FPT) increases as the cell volume at birth V_0 is increased. Using the expression of the first-passage time in (5.11), the noise (CV^2) is numerically computed for different values of the newborn cell volume. The parameters used for the model are $k_m = 1$ (1/min), $\theta = 100$ (molecules), $\alpha = 0.03$ (1/min), and geometric distribution of the burst size $\boldsymbol{B}$ with mean $b = 3$ (molecules). *Right*: The prediction from the model is validated using the experimental data from [4]. Using bootstrapping, a statistically significant (p-value=0.00023) increase in the noise is observed from smaller initial volume to a larger initial volume.

(see Fig. 5.3 (left)). This prediction is consistent with data from [4], as shown on right part of Fig. 5.3 . The noise behavior can be explained by observing that on average a cell with smaller volume at birth takes more time for division. Therefore the fluctuations are time averaged, leading to a smaller noise in division time as compared to that of a cell with larger volume at birth.

5.3 Distribution of the Volume Added Between Divisions

In the previous section, we determined the distribution of the division time for a cell of given initial volume V_0. Combining this with the fact that volume of the cell grows exponentially in time, we can obtain the distribution of the volume added to V_0. More precisely, the volume added from birth to division (denoted by $\boldsymbol{\Delta}$) is given as

$$\boldsymbol{\Delta} = V_0 \left(e^{\alpha \mathbf{T}} - 1 \right). \tag{5.12}$$

The distribution of FPT given in (5.11) can be used to find the distribution of $\mathbf{\Delta}$ as follows.

$$\mathbb{P}\{\mathbf{\Delta} \leq v\} = Prob\left\{V_0\left(e^{\alpha \mathbf{T}} - 1\right) \leq v\right\} \tag{5.13}$$

$$= Prob\left\{\mathbf{T} \leq \frac{1}{\alpha}\ln\left(\frac{v}{V_0}+1\right)\right\} \tag{5.14}$$

$$= \int_0^{\frac{1}{\alpha}\ln\left(\frac{v}{V_0}+1\right)} \rho_{\mathbf{T}}(t)dt. \tag{5.15}$$

Therefore the probability density function of $\mathbf{\Delta}$ is given by

$$\rho_{\mathbf{\Delta}}(v) = \frac{d}{dv}\left(\mathbb{P}\{\mathbf{\Delta} \leq v\}\right) \tag{5.16}$$

$$= \frac{d}{dv}\int_0^{\frac{1}{\alpha}\ln\left(\frac{v}{V_0}+1\right)} \rho_{\mathbf{T}}(t)dt \tag{5.17}$$

$$= \rho_{\mathbf{T}|V_0}\left(\frac{1}{\alpha}\ln\left(\frac{v}{V_0}+1\right)\right)\frac{d}{dv}\left(\frac{1}{\alpha}\ln\left(\frac{v}{V_0}+1\right)\right). \tag{5.18}$$

Note that $\dfrac{d}{dv}\left(\dfrac{1}{\alpha}\ln\left(\dfrac{v}{V_0}+1\right)\right) = \dfrac{1}{\alpha\left(V_0+v\right)}$. Also

$$R\left(\frac{1}{\alpha}\ln\left(\frac{v}{V_0}+1\right)\right) = \frac{k_m v}{\alpha}, \quad r\left(\frac{1}{\alpha}\ln\left(\frac{v}{V_0}+1\right)\right) = k_m\left(v+V_0\right). \tag{5.19}$$

Hence we can write the probability density of $\mathbf{\Delta}$ as

$$\rho_{\mathbf{\Delta}}(v) = \frac{1}{\alpha\left(V_0+v\right)}\sum_{n=1}^{\infty}\frac{\left(\frac{k_m v}{\alpha}\right)^{n-1}}{(n-1)!}k_m\left(v+V_0\right)\exp\left(-\frac{k_m v}{\alpha}\right)\rho_{\mathbf{N}}(n) \tag{5.20}$$

$$= \sum_{n=1}^{\infty}\frac{\left(\frac{k_m v}{\alpha}\right)^{n-1}}{(n-1)!}\frac{k_m}{\alpha}\exp\left(-\frac{k_m v}{\alpha}\right)\rho_{\mathbf{N}}(n). \tag{5.21}$$

Notice that this distribution is an Erlang distribution conditioned to the random variable $\mathbf{N}$.

One striking observation is that $\rho_{\mathbf{\Delta}}(v)$ is independent of the initial volume V_0 (see Fig. 5.1(d) for a depiction of this). This is in agreement with the experimental observations that the distribution of the added volume does not depend on the size of the cell at birth [4]. Further, we can also use the probability density function to find moments of $\mathbf{\Delta}$. We will first discuss the expression of mean $\mathbf{\Delta}$, particularly emphasizing its dependence on the growth rate. The higher order moments are taken up next.

5.3.1 Mean of Added Volume

The distribution of $\mathbf{\Delta}$ given in (5.21) is an Erlang distribution conditioned on N, the minimum number of burst events required for protein level $x(t)$ to cross a threshold θ. The formula for $\langle \mathbf{\Delta} \rangle$ can be written as

$$\langle \mathbf{\Delta} \rangle = \frac{\alpha}{k_m} \langle \mathbf{N} \rangle = \frac{\alpha}{k_m} \left(\frac{\theta}{b} + 1 \right), \tag{5.22}$$

where the expression of $\langle \mathbf{N} \rangle$ is obtained by assuming that protein is produced in geometric bursts with mean burst size b [171]. The formula in (5.22) shows that If k_m is independent of growth rate α, the average volume added $\mathbf{\Delta}$ is a linear function of α which is in agreement with experimental data on *Pseudomonas aeruginosa* [199]. Interestingly, it has been seen that $\mathbf{\Delta}$ can depend on α in several different ways. For instance, *Caulobacter crescentus* exhibits an added volume independent of α whereas this relationship is thought to be exponential in case of *Escherichia coli* [4, 198]. One way by which our model can result in $\mathbf{\Delta}$ being independent of α could be considering k_m a linear function of α. Next we discuss an approach by which an exponential dependency can be seen.

So far we have considered that the time-keeper protein observes time between birth to division which, in principle, can be between other important events in the cell cycle as well. Here we consider that the initiation of DNA replication takes place when sufficient time-keeper protein has been accumulated per origin of replication [82, 204]. The corresponding division event is assumed to occur with a constant delay of T after an initiation. The delay τ_d here is what is called "$C + D$" period whereby C represents the time to replicate the DNA and D denotes the time between DNA replication and division [221, 222]. In this formulation, the volume added between two consecutive initiation events for each origin of replication is same as $\mathbf{\Delta}$ in our previous model. Further, the average volume added between divisions associated with these initiations (denoted by $\mathbf{\Delta}^*$) is related with $\mathbf{\Delta}$ as [206]

$$\langle \mathbf{\Delta}^* \rangle \approx \langle \mathbf{\Delta} \rangle \, e^{\alpha \tau_d} = \frac{\alpha}{k_m} \left(\frac{\theta}{b} + 1 \right) e^{\alpha \tau_d}. \tag{5.23}$$

When the parameter k_m is constant with respect to the growth rate α, the expression in (5.23) suggests two different regimes of how $\mathbf{\Delta}^*$ depends upon α. For small values of α, we have $\alpha \exp(\alpha\tau_d) \approx \alpha$, i.e., the mean added volume increases linearly with the growth rate. In the regime where α is large, the exponential term dominates giving rise to exponential dependence. This also means that if the growth rate α is small, it is not possible to distinguish whether the underlying mechanism accounts for volume added between two division events or two initiation events as the data will show linear dependence of the average added volume with changes in α [199]. Notice that a pure exponential relationship between $\mathbf{\Delta}^*$ and α can also be obtained by having k_m a linearly increasing function of α. For this particular case, the volume accounted by each origin of replication $\mathbf{\Delta}$ will become invariant of the growth rate, consistent with analysis by [206], and [223].

To sum up, how the volume added between birth to division depends upon the growth rate differs between species. A variety of relationships can be obtained depending upon how k_m varies with growth rate α, and whether the time-keeper protein accounts for two division events or two other events in the cell cycle. We now go back to discussing the higher order moments of the division to division model.

5.3.2 Higher Order Moments of Added Volume and Scale Invariance of Distributions

We can use the distribution of $\mathbf{\Delta}$ to get insight into its higher order statistics such as coefficient of variation squared ($CV^2_{\mathbf{\Delta}}$), skewness ($skew_{\mathbf{\Delta}}$), etc. For example, when the protein production is considered in geometric bursts, we have the following

$$CV^2_{\mathbf{\Delta}} = \frac{b^2 + 2b\theta + \theta}{(b+\theta)^2}, \quad skew_{\mathbf{\Delta}} = \frac{2\left(b^3 + 3b^2\theta + 3b\theta + \theta\right)}{\left(b^2 + 2b\theta + \theta\right)^{3/2}}. \tag{5.24}$$

It can be noticed that skewness is a positive quantity, consistent with previous understanding [197]. Moreover, both CV^2 and skewness are independent of the growth rate α. It turns out an even more general property is true: an appropriately scaled j^{th} order moment of $\mathbf{\Delta}$, i.e., $\left\langle \mathbf{\Delta}^j \right\rangle / \left\langle \mathbf{\Delta} \right\rangle^j$ is independent of α, notwithstanding the underlying

distribution of burst size. This arises from the fact that the distribution of $\mathbf{\Delta}$ can be written in the following form

$$\rho_{\mathbf{\Delta}}(v) = \frac{1}{\langle \mathbf{\Delta} \rangle} \overline{\rho}_{\mathbf{\Delta}} \left(\frac{v}{\langle \mathbf{\Delta} \rangle} \right), \tag{5.25}$$

for any distribution of burst size. An important implication of above form of distribution of $\mathbf{\Delta}$ is the scale invariance property: the shape of the distribution across different growth rates is essentially same, and a single parameter $\langle \mathbf{\Delta} \rangle$ is sufficient to characterize the distribution of $\mathbf{\Delta}$ [183]. Recent experimental data have also exhibited the scale invariance property [4, 203, 224].

Interestingly, the above invariance property is not limited to the distribution of the added volume $\mathbf{\Delta}$. Ignoring the partitioning errors in the volume, it can be seen that in steady-state the cell-size distribution at birth is approximately same as the distribution of $\mathbf{\Delta}$ [4]. Also, the size at division is $2\mathbf{\Delta}$. Thus, the scale invariance of $\mathbf{\Delta}$ immediately implies scale invariance of the distributions of cell sizes at birth and division [4]. Moreover, the distribution of the division time can also be determined by unconditioning (5.11) with respect to the distribution of the initial volume V_0. It can be also seen that the distribution of the division time also has the scale invariance property which is in agreement with the results in [210, 225].

5.4 Biological Relevance and Discussion

It is now well understood that several prokaryotes employ an adder mechanism to achieve cell size homeostasis wherein on average a constant volume is added between birth to division [4, 198–200]. In this work, we studied a molecular mechanism that can result in the adder principle of cell size control. This mechanism considers a protein sensing the volume added between birth to division or two other events in the cell cycle [82, 199, 201–206, 226]. Our work shows that in addition to the prior understanding that this biophysical mechanism manifests as adder principle in the mean sense, it can also explain the stochastic attributes of various quantities such as cell-division time, size added between birth to division, etc. In particular, we get the following important results:

- The distribution of volume added between birth to division is independent of the cell volume at birth. This is consistent with observations by [4].

- The distributions of key quantities such as the added volume, division time, volume at birth, volume at division, etc. show the scale invariance property [4, 183, 210].
- The noise in division time increases with increase in cell size at birth, which was validated from the available data from [4].

In addition, we also discussed how the volume added between birth and division changes as the growth rate is varied. Our model produces a wide range of behaviors such as linear, exponential or no dependency of the added volume on the growth rate. Next, we discuss the implications of the findings of this paper.

5.4.1 Potential Candidates for the Time–Keeper Protein

Among many proteins involved in the process, prominent candidates for the time-keeper are FtsZ and DnaA. More specifically if the constant volume addition is considered between division to division, FtsZ is thought to act as the time-keeper protein [227–230]. It has been proposed that the accumulation of FtsZ up to a critical level is required for cell division [231–233]. Likewise, the protein DnaA is known to regulate the timing of initiation of replication, if the constant volume is added between two initiation events [230, 234, 235]. In this case, initiation is proposed to occur when a critical number (around 20) of DnaA-ATP molecules are available [236]. While it is not clear yet whether the production of DnaA or its conversion to DnaA-ATP is a rate limiting step in the initiation process, the model presented here can account for both cases as long as the conversion to DnaA-ATP happens at a volume dependent rate.

We can employ the closed-form expressions for the moments of $\mathbf{\Delta}$ developed in this work to investigate roles of these candidate proteins. We illustrate this by considering geometrically distributed burst of proteins for which the expressions of mean and coefficient of variation squared (CV^2) are given by (5.22) and (5.24) respectively. Thus, increasing the threshold θ or decreasing the mean burst size of the time-keeper protein b should result in decrease in the CV^2 of the added volume. A commensurate change in k_m by making the promoter weak or stronger can be further used to ensure a fixed

mean $\mathbf{\Delta}$. Experimentally, the mean burst size can be altered by changing the translation rate of the proteins using techniques such as mutations in the Shine-Dalgarno sequence. Changing the threshold can be achieved by changing the protein sequence which affects its function and thus leads to a different number of protein molecules being required for division.

5.4.2 Deviations from the Adder Principle

Recently it has been proposed that cells employ a generalized version of the adder principle wherein the volume added between divisions $\mathbf{\Delta}$ depends upon the cell volume at birth V_0 [237, 238]. Intuitively, mechanisms which would lead to a smaller (longer) division time than the adder principle would result in a decrease (increase) in $\mathbf{\Delta}$ as V_0 is increased. There could be several ways to realize these deviations from the adder principle in our model. For instance, if we consider that the time-keeper protein does not degrade fully upon division and the remaining proteins are divided in the daughter cells in proportion to their respective volumes at birth, the added volume decreases as volume of daughter cell is increased. This is because if there are already time-keeper proteins present in the cell at the time of its birth, the threshold will be achieved earlier than the case when there were no proteins at birth. Another possible way of getting such deviation could be if the mean burst size is an increasing function of the cell volume which leads to a smaller time to reach the copy number threshold of the protein. In contrast, $\mathbf{\Delta}$ being an increasing function of V_0 can be obtained by curbing the scaling of protein accumulation with the cell volume. One possibility is to assume a transcription rate of the form $r(t) = k_m \frac{V(t)}{V(t)+\overline{V}}$ where the transcription rate saturates with increase in volume. Alternatively, this effect could also be achieved by considering that instead of cell division occurring upon achieving a constant volume addition, its propensity increases as the added volume increases.

BIBLIOGRAPHY

[1] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–434, 1976.

[2] Niraj Kumar, Thierry Platini, and Rahul V. Kulkarni. Exact Distributions for Stochastic Gene Expression Models with Bursting and Feedback. *Physical Review Letters*, 113:268105, 2014.

[3] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81:2340–2361, 1977.

[4] Sattar Taheri-Araghi, Serena Bradde, John T. Sauls, Norbert S. Hill, Petra Anne Levin, Johan Paulsson, Massimo Vergassola, and Suckjoon Jun. Cell-size control and homeostasis in bacteria. *Current Biology*, 25:385–391, 2015.

[5] Arjun Raj and Alexander van Oudenaarden. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135:216–226, 2008.

[6] J. M. Raser and E. K. O'Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309:2010–2013, 2005.

[7] Mads Kærn, Timothy C Elston, William J Blake, and James J Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, 6:451–464, 2005.

[8] Avigdor Eldar and Michael B Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467:167–173, 2010.

[9] Gregor Neuert, Brian Munsky, Rui Zhen Tan, Leonid Teytelman, Mustafa Khammash, and Alexander van Oudenaarden. Systematic identification of signal-activated stochastic gene regulation. *Science*, 339:584–587, 2013.

[10] Guilhem Chalancon, Charles N.J. Ravarani, S. Balaji, Alfonso Martinez-Arias, L. Aravind, Raja Jothi, and Madan Babu. Interplay between gene expression noise and regulatory network architecture. *Trends in Genetics*, 28:221–232, 2012.

[11] Angeliki Magklara and Stavros Lomvardas. Stochastic gene expression in mammals: lessons from olfaction. *Trends in Cell Biology*, 23:449–456, 2014.

[12] Gürol M Süel, Jordi Garcia-Ojalvo, Louisa M Liberman, and Michael B Elowitz. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature*, 440:545–550, 2006.

[13] Hédia Maamar, Arjun Raj, and David Dubnau. Noise in gene expression determines cell fate in bacillus subtilis. *Science*, 317:526–529, 2007.

[14] Roy D. Dar, Nina N. Hosmane, Michelle R. Arkin, Robert F. Siliciano, and Leor S. Weinberger. Screening for noise in gene expression identifies drug synergies. *Science*, 344:1392–1396, 2014.

[15] Nico Battich, Thomas Stoeger, and Lucas Pelkmans. Control of transcript variability in single mammalian cells. *Cell*, 163:1596–1610, 2015.

[16] Iain G. Johnston, Bernadett Gaal, Ricardo Pires das Neves, Tariq Enver, Francisco J. Iborra, and Nick S. Jones. Mitochondrial variability as a source of extrinsic cellular noise. *PLOS Computational Biology*, 8:e1002416, 2012.

[17] A. P. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics*, 149:1633–1648, 1998.

[18] Richard Losick and Claude Desplan. Stochasticity and cell fate. *Science*, 320:65–68, 2008.

[19] Gábor Balázsi, Alexander van Oudenaarden, and James J Collins. Cellular decision making and biological noise: From microbes to mammals. *Cell*, 144:910–925, 2014.

[20] Thomas M Norman, Nathan D Lord, Johan Paulsson, and Richard Losick. Memory and modularity in cell-fate decision making. *Nature*, 503:481–486, 2013.

[21] Francois St-Pierre and Drew Endy. Determination of cell fate selection during phage lambda infection. *Proceedings of the National Academy of Sciences*, 105:20705–20710, 2008.

[22] Kyung H. Kim and Herbert M. Sauro. Adjusting phenotypes by noise control. *PLOS Computational Biology*, 8:e1002344, 2012.

[23] H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, and S. Huang. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*, 453:544–547, 2008.

[24] Elsa Abranches, Ana M. V. Guedes, Martin Moravec, Hedia Maamar, Petr Svoboda, Arjun Raj, and Domingos Henrique. Stochastic nanog fluctuations allow mouse embryonic stem cells to explore pluripotency. *Development*, 141:2770–2779, 2014.

[25] Maria Elena Torres-Padilla and Ian Chambers. Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. *Development*, 141:2173–2181, 2014.

[26] R. L. Thompson, C. M. Preston, and N. M. Sawtell. De novo synthesis of VP16 coordinates the exit from HSV latency in vivo. *PLOS Pathogens*, 5:e1000352, 2009.

[27] Murat Acar, Jerome T. Mettetal, and Alexander van Oudenaarden. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics*, 40:471–475, 2008.

[28] Jan-Willem Veening, Wiep Klass Smits, and Oscar P Kuipers. Bistability, epigenetics, and bet-hedging in bacteria. *Annual Review of Microbiology*, 62:193–210, 2008.

[29] E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309:2075–2078, 2005.

[30] Amy L. Bishop, Faiza A. Rab, Edward R. Sumner, and Simon V. Avery. Phenotypic heterogeneity can enhance rare-cell survival in stress-sensitive yeast populations. *Molecular Microbiology*, 63:507–520, 2007.

[31] Martin Ackermann. A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Reviews Microbiology*, 13:497–508, 2015.

[32] Che-Chi Shu, Anushree Chatterjee, Wei-Shou Hu, and Doraiswami Ramkrishna. Role of intracellular stochasticity in biofilm growth. insights from population balance modeling. *PLOS ONE*, 8:e79196, 2013.

[33] E. Libby, T. J. Perkins, and P. S. Swain. Noisy information processing through transcriptional regulation. *Proceedings of the National Academy of Sciences*, 104:7151–7156, 2007.

[34] H. B. Fraser, A. E. Hirsh, G. Giaever, J. Kumm, and M. B. Eisen. Noise minimization in eukaryotic gene expression. *PLOS Biology*, 2:e137, 2004.

[35] B. Lehner. Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Molecular Systems Biology*, 4:170, 2008.

[36] R. Kemkemer, S. Schrank, W. Vogel, H. Gruler, and D. Kaufmann. Increased noise as an effect of haploinsufficiency of the tumor-suppressor gene neurofibromatosis type 1 in vitro. *Proceedings of the National Academy of Sciences*, 99:13783–13788, 2002.

[37] D. L. Cook, A. N. Gerber, and S. J. Tapscott. Modeling stochastic gene expression: implications for haploinsufficiency. *Proceedings of the National Academy of Sciences*, 95:15641–15646, 1998.

[38] R. Bahar, C. H. Hartmann, K. A. Rodriguez, A. D. Denny, R. A. Busuttil, M. E. Dolle, R. B. Calder, G. B. Chisholm, B. H. Pollock, C. A. Klein, and J. Vijg. Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature*, 441:1011–1014, 2006.

[39] Amy Brock, Hannah Chang, and Sui Huang. Non-genetic heterogeneity – a mutation-independent driving force for the somatic evolution of tumours. *Nature Reviews Genetics*, 10:336–342, 2009.

[40] N.Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305:1622–1625, 2004.

[41] Jingchen Feng, David A. Kessler, Eshel Ben-Jacob, and Herbert Levine. Growth feedback as a basis for persister bistability. *Proceedings of the National Academy of Sciences*, 111:544–549, 2014.

[42] Etienne Maisonneuve, Manuela Castro-Camargo, and Kenn Gerdes. (p)ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. *Cell*, 154:1140–1150, 2013.

[43] Imane El Meouche, Yik Siu, and Mary J. Dunlop. Stochastic expression of a multiple antibiotic resistance activator confers transient resistance in single cells. *Scientific Reports*, 6:19538, 2016.

[44] Iftach Nachman, Aviv Regev, and Sharad Ramanathan. Dissecting timing variability in yeast meiosis. *Cell*, 131:544–556, 2007.

[45] Amnon Amir, Oren Kobiler, Assaf Rokney, Amos B Oppenheim, and Joel Stavans. Noise in timing and precision of gene activities in a genetic cascade. *Molecular Systems Biology*, 3:71, 2007.

[46] Fabian Fröhlich, Philipp Thomas, Atefeh Kazeroonian, Fabian J Theis, Ramon Grima, and Jan Hasenauer. Inference for stochastic chemical kinetics using moment equations and system size expansion. *PLoS Comput Biol*, 12:e1005030, 2016.

[47] Brian Munsky, Zachary Fox, and Gregor Neuert. Integrating single-molecule experiments and discrete stochastic models to understand heterogeneous gene transcription dynamics. *Methods*, 85:12–21, 2015.

[48] P. Milner, C. Gillespie, and D. Wilkinson. Moment closure based parameter inference of stochastic kinetic models. *Statistics and Computing*, pages 1–9, 2012.

[49] Christoph Zechner, Jakob Ruess, Peter Krenn, Serge Pelet, Matthias Peter, John Lygeros, and Heinz Koeppl. Moment-based inference predicts bimodality in transient gene expression. *Proceedings of the National Academy of Sciences*, 109:8340–8345, 2012.

[50] Philipp Kügler. Moment fitting for parameter inference in repeatedly and partially observed stochastic biological models. *PLOS ONE*, 7:e43001, 2012.

[51] I. Nasell. Moment closure and the stochastic logistic model. *Theoretical Population Biology*, 63:159–168, 2003.

[52] I. Nasell. An extension of the moment closure method. *Theoretical Population Biology*, 64:233–239, 2003.

[53] T. J. Newman, Jean-Baptiste Ferdy, and C. Quince. Extinction times and moment closure in the stochastic logistic process. *Journal of Theoretical Biology*, 65:115–126, 2004.

[54] I. Krishnarajah, A. Cook, G. Marion, and G. Gibson. Novel moment closure approximations in stochastic epidemics. *Bulletin of Mathematical Biology*, 67:855–873, 2005.

[55] A. Singh and J. P. Hespanha. Moment closure techniques for stochastic models in population biology. In *Proc. of the 2006 Amer. Control Conference, Minneapolis, MN*, 2006.

[56] A. Singh and J. P. Hespanha. Lognormal moment closures for biochemical reactions. In *Proc. of the 45th IEEE Conf. on Decision and Control, San Diego*, 2006.

[57] A. Singh and J. P. Hespanha. Stochastic analysis of gene regulatory networks using moment closure. In *Proc. of the 2007 Amer. Control Conference, New York, NY*, 2007.

[58] C. H. Lee, K. Kim, and P. Kim. A moment closure method for stochastic reaction networks. *Journal of Chemical Physics*, 130:134107, 2009.

[59] A. Singh and J. P. Hespanha. A derivative matching approach to moment closure for the stochastic logistic model. *Bulletin of Mathematical Biology*, 69:1909–1925, 2007.

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

[61] Ramon Grima. A study of the accuracy of moment-closure approximations for stochastic chemical kinetics. *Journal of Chemical Physics*, 136:154105, 2012.

[62] Patrick Smadbeck and Yiannis N. Kaznessis. A closure scheme for chemical master equations. *Proceedings of the National Academy of Sciences*, 110:14261–14265, 2013.

[63] Mahommad Soltani, Cesar A. Vargas-Garcia, and Abhyudai Singh. Conditional moment closure schemes for studying stochastic dynamics of genetic circuits. *IEEE Transactions on Biomedical Systems and Circuits*, 9:518–526, 2015.

[64] Peter Milner, Colin S. Gillespie, and Darren J. Wilkinson. Moment closure approximations for stochastic kinetic models with rational rate laws. *Mathematical Biosciences*, 231:99–104, 2011.

[65] Cheng Ly and Daniel Tranchina. Critical analysis of dimension reduction by a moment closure method in a population density approach to neural network modeling. *Neural Computation*, 19:2032–2092, 2007.

[66] Jiangmeng Zhang, L. DeVille, S. Dhople, and A.D. Dominguez-Garcia. A maximum entropy approach to the moment closure problem for stochastic hybrid systems at equilibrium. In *Proc. of the 53rd IEEE Conf. on Decision and Control, Los Angeles, CA*, pages 747–752, 2014.

[67] D. Schnoerr, G. Sanguinetti, and R. Grima. Validity conditions for moment closure approximations in stochastic chemical kinetics. *The Journal of Chemical Physicsl*, 141:084103, 2014.

[68] J. Hasenauer, V. Wolf, A. Kazeroonian, and F.J. Theis. Method of conditional moments (mcm) for the chemical master equation. *Journal of Mathematical Biology*, 69:687–735, 2014.

[69] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Comparison of different moment-closure approximations for stochastic chemical kinetics. *The Journal of Chemical Physics*, 143:185101, 2015.

[70] Christian Kuehn. Moment closure–a brief review. In Eckehard Schöll, Sabine H. L. Klapp, and Philipp Hövel, editors, *Control of Self-Organizing Nonlinear Systems*, Understanding Complex Systems, pages 253–271. Springer International Publishing, 2016.

[71] Guillaume Lambert and Edo Kussell. Quantifying selective pressures driving bacterial evolution using lineage analysis. *Physical Review X*, 5:011016, 2015.

[72] Sergiy Bogomolov, Thomas A. Henzinger, Andreas Podelski, Jakob Ruess, and Christian Schilling. Adaptive Moment Closure for Parameter Inference of Biochemical Reaction Networks. In Olivier Roux and Jérémie Bourdon, editors, *Computational Methods in Systems Biology*, Lecture Notes in Computer Science, pages 77–89. Springer International Publishing, 2015.

[73] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Approximation and inference methods for stochastic biochemical kinetics–a tutorial review. *Journal of Physics A: Mathematical and Theoretical*, 50:093001, 2017.

[74] HK Salz. Male or female? The answer depends on when you ask. *PLOS Biology*, 5:e335, 2007.

[75] Yifat Goldschmidt, Evgeny Yurkovsky, Amit Reif, Roni Rosner, Amit Akiva, and Iftach Nachman. Control of relative timing and stoichiometry by a master regulator. *PLOS ONE*, 10:1–14, 2015.

[76] H. H. McAdams and A. P. Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94:814–819, 1997.

[77] Juan M Pedraza and Johan Paulsson. Random timing in signaling cascades. *Molecular Systems Biology*, 3:81, 2007.

[78] Katherine C Chen, Laurence Calzone, Attila Csikasz-Nagy, Frederick R Cross, Bela Novak, and John J Tyson. Integrative analysis of cell cycle control in budding yeast. *Molecular Biology of the Cell*, 15:3841–3862, 2004.

[79] James M Bean, Eric D Siggia, and Frederick R Cross. Coherence and timing of cell cycle start examined at single-cell resolution. *Molecular Cell*, 21:3–14, 2006.

[80] David L Satinover, David L Brautigan, and P Todd Stukenberg. Aurora-A kinase and inhibitor-2 regulate the cyclin threshold for mitotic entry in xenopus early embryonic cell cycles. *Cell Cycle*, 5:2268–2274, 2006.

[81] Xili Liu, Xin Wang, Xiaojing Yang, Sen Liu, Lingli Jiang, Yimiao Qu, Lufeng Hu, Qi Ouyang, and Chao Tang. Reliable cell cycle commitment in budding yeast is ensured by signal integration. *eLife*, 4:e03977, 2015.

[82] Ariel Amir. Cell size regulation in bacteria. *Physical Review Letters*, 112:208102, 2014.

[83] David T Champlin and James W Truman. Ecdysteroid control of cell proliferation during optic lobe neurogenesis in the moth manduca sexta. *Development*, 125:269–277, 1998.

[84] Takashi Koyama, Masafumi Iwami, and Sho Sakurai. Ecdysteroid control of cell cycle and cellular commitment in insect wing imaginal discs. *Molecular and Cellular Endocrinology*, 213:155–166, 2004.

[85] Karen Carniol, Patrick Eichenberger, and Richard Losick. A threshold mechanism governing activation of the developmental regulatory protein σf in bacillus subtilis. *Journal of Biological Chemistry*, 279:14860–14870, 2004.

[86] Patrick J Piggot and David W Hilbert. Sporulation of bacillus subtilis. *Current Opinion in Microbiology*, 7:579–586, 2004.

[87] Sabrina L Spencer, Suzanne Gaudet, John G Albeck, John M Burke, and Peter K Sorger. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, 459:428–432, 2009.

[88] Jérémie Roux, Marc Hafner, Samuel Bandara, Joshua J Sims, Hannah Hudson, Diana Chai, and Peter K Sorger. Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. *Molecular Systems Biology*, 11:803, 2015.

[89] M Kracikova, G Akiri, A George, R Sachidanandam, and SA Aaronson. A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis. *Cell Death & Differentiation*, 20:576–588, 2013.

[90] Khem Raj Ghusinga, Cesar A Vargas-Garcia, Andrew Lamperski, and Abhyudai Singh. Exact lower and upper bounds on stationary moments in stochastic biochemical systems. *Physical Biology*, 14:04LT01, 2017.

[91] Andrew Lamperski, Khem Raj Ghusinga, and Abhyudai Singh. Analysis and control of stochastic systems using semidefinite programming over moments. *IEEE Transactions on Automatic Control (to appear)*, 2019.

[92] Khem Raj Ghusinga, Andrew Lamperski, and Abhyudai Singh. Moment analysis of stochastic hybrid systems using semidefinite programming. *arXiv preprint arXiv:1802.00376*, 2018.

[93] Andrew Lamperski, Khem Raj Ghusinga, and Abhyudai Singh. Stochastic optimal control using semidefinite programming for moment dynamics. *Proc. of the 55th IEEE Conf. on Decision and Control, Las Vegas*, pages 1990–1995, 2016.

[94] Khem Raj Ghusinga, Andrew Lamperski, and Abhyudai Singh. Estimating stationary characteristic functions of stochastic systems via semidefinite programming. *European Control Conference*, 2018.

[95] Khem Raj Ghusinga, John J. Dennehy, and Abhyudai Singh. First-passage time approach to controlling noise in the timing of intracellular events. *Proceedings of the National Academy of Sciences*, 114:693–698, 2017.

[96] Khem Raj Ghusinga and Abhyudai Singh. Effect of gene-expression bursts on stochastic timing of cellular events. *Proc. of the 2017 Amer. Control Conference, Seattle, WA*, pages 2118–2123, 2017.

[97] Khem Raj Ghusinga and Abhyudai Singh. Optimal regulation of protein degradation to schedule cellular events with precision. *Proc. of the 2016 Amer. Control Conference, Boston, MA*, 2016.

[98] Khem Raj Ghusinga, Pak-Wing Fok, and Abhyudai Singh. Optimal auto–regulation to minimize first–passage time variability in protein level. *Proc. of the 2015 Amer. Control Conference, Chicago, IL*, pages 4011–4016, 2015.

[99] K. R. Ghusinga and A. Singh. Theoretical predictions on the first-passage time for a gene expression model. *Proc. of the 54^{th} IEEE Conf. on Decision and Control, Osaka, Japan*, pages 3864–3869, 2015.

[100] K. R. Ghusinga and A. Singh. First-passage time calculations for a gene expression model. *Proc. of the 53rd IEEE Conf. on Decision and Control, Los Angeles, CA*, pages 3047–3052, 2014.

[101] Khem Raj Ghusinga, Cesar A. Vargas-Garcia, and Abhyudai Singh. A mechanistic stochastic framework for regulating bacterial cell division. *Scientific Reports*, page 30229, 2016.

[102] D. A. McQuarrie. Stochastic approach to chemical kinetics. *Journal of Applied Probability*, 4:413–478, 1967.

[103] NG Van Kampen. *Stochastic processes in physics and chemistry*. Elsevier, 2011.

[104] Crispin W Gardiner et al. *Handbook of stochastic methods*, volume 4. Springer Berlin, 1985.

[105] D. T. Gillespie and L. R. Petzold. Improved leap-size selection for accelerated stochastic simulation. *Journal of Chemical Physics*, 119:8229–8234, 2003.

[106] Muruhan Rathinam, Linda R. Petzold, Yang Cao, and Daniel T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *Journal of Chemical Physics*, 119:12784–12794, 2003.

[107] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics*, 115:1716–1733, 2001.

[108] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A*, 104:1876–1889, 2000.

[109] Y. Cao, H. Li, and L. Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *Journal of Chemical Physics*, 121:4059–4067, 2004.

[110] B. Munsky and M. Khammash. The finite state projection algorithm for the solution of the chemical master equation. *Journal of Chemical Physics*, 124:044104, 2006.

[111] D. F. Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *Journal of Chemical Physics*, 127:214107, 2007.

[112] Jifeng Hu, Hye Kang, and Hans G. Othmer. Stochastic analysis of reaction-diffusion processes. *Bulletin of Mathematical Biology*, 76:854–894, 2014.

[113] Youfang Cao, Anna Terebus, and Jie Liang. State space truncation with quantified errors for accurate solutions to discrete chemical master equation. *Bulletin of Mathematical Biology*, 78:617–661, 2016.

[114] Youfang Cao, Anna Terebus, and Jie Liang. Accurate chemical master equation solution using multi-finite buffers. *Multiscale Modeling & Simulation*, 14:923–963, 2016.

[115] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Series in Electrical Engineering. McGraw-Hill, Inc., New York, 3rd edition, 1991.

[116] Alexander Andreychenko, Linar Mikeev, and Verena Wolf. Model reconstruction for moment-based stochastic chemical kinetics. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 25:12, 2015.

[117] J. P. Hespanha and A. Singh. Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems. *International Journal of Robust and Nonlinear Control*, 15:669–689, 2005.

[118] A. Singh and J. P. Hespanha. Approximate moment dynamics for chemically reacting systems. *IEEE Transactions on Automatic Control*, 56:414–418, 2011.

[119] A. Singh and J. P. Hespanha. Stochastic hybrid systems for studying biochemical processes. *Philosophical Transactions of the Royal Society A*, 368:4995–5011, 2010.

[120] E. Sontag and A. Singh. Exact moment dynamics for feedforward nonlinear chemical reaction networks. *IEEE Life Sciences Letters*, 1:26–29, 2015.

[121] Jean Bernard Lasserre. *Moments, positive polynomials and their applications*, volume 1. World Scientific, 2009.

[122] Carl D Meyer. *Matrix analysis and applied linear algebra*. Siam, 2000.

[123] Sean P Meyn and Richard L Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.

[124] Lee DeVille, Sairaj Dhople, Alejandro D Domínguez-García, and Jiangmeng Zhang. Moment closure and finite-time blowup for piecewise deterministic markov processes. *SIAM Journal on Applied Dynamical Systems*, 15(1):526–556, 2016.

[125] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[126] Linda JS Allen. *An introduction to stochastic processes with applications to biology*. Prentice Hall, 2003.

[127] J. H. Matis and T. R. Kiffe. Effects of immigration on some stochastic logistic models: A cumulant truncation analysis. *Theoretical Population Biology*, 56:139–161, 1999.

[128] I. Nasell. Extinction and quasi-stationarity in the verhulst logistic model. *Journal of Theoretical Biology*, 211:11–27, 2001.

[129] A. Singh and J. P. Hespanha. Evolution of autoregulation in the presence of noise. *IET Systems Biology*, 3:368–378, 2009.

[130] M Soltani, C A Vargas-Garcia, N Kumar, R Kulkarni, and A Singh. Approximate statistical dynamics of a genetic feedback circuit. *Proc. of the 2015 Amer. Control Conference, Chicago, IL*, pages 4424–4429, 2015.

[131] Stephan Bohacek, Joao P Hespanha, Junsoo Lee, and Katia Obraczka. A hybrid systems modeling framework for fast and accurate simulation of data communication networks. In *ACM SIGMETRICS Performance Evaluation Review*, volume 31, pages 58–69. ACM, 2003.

[132] João P Hespanha. Stochastic hybrid systems: Application to communication networks. In *Hybrid systems: computation and control*, pages 387–401. Springer, 2004.

[133] Joao P Hespanha. A model for stochastic hybrid systems with application to communication networks. *Nonlinear Analysis: Theory, Methods & Applications*, 62(8):1353–1383, 2005.

[134] Jianghai Hu. Application of stochastic hybrid systems in power management of streaming data. In *American Control Conference, 2006*, pages 6–pp. IEEE, 2006.

[135] Duarte Antunes, Joao P Hespanha, and Carlos Silvestre. Stochastic hybrid systems with renewal transitions: Moment analysis with application to networked control systems with delays. *SIAM Journal on Control and Optimization*, 51(2):1481–1499, 2013.

[136] João P. Hespanha. Modeling and analysis of networked control systems using stochastic hybrid systems. *Annual Reviews in Control*, 38(2):155–170, 2014.

[137] Evelyn Buckwar and Martin G Riedler. An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *Journal of mathematical biology*, 63(6):1051–1093, 2011.

[138] Andrea Lecchini Visintini, William Glover, John Lygeros, and Jan Maciejowski. Monte carlo optimization for conflict resolution in air traffic control. *IEEE Transactions on Intelligent Transportation Systems*, 7(4):470–482, 2006.

[139] Weiyi Liu and Inseok Hwang. Probabilistic trajectory prediction and conflict detection for air traffic control. *Journal of Guidance, Control and Dynamics*, 34(6):1779–1789, 2011.

[140] Jianghai Hu and Maria Prandini. Aircraft conflict detection: a method for computing the probability of conflict based on markov chain approximation. In *European Control Conference (ECC), 2003*, pages 2225–2230. IEEE, 2003.

[141] Martin Střelec, Karel Macek, and Alessandro Abate. Modeling and simulation of a microgrid as a stochastic hybrid system. In *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pages 1–9. IEEE, 2012.

[142] Alexandre David, Kim G Larsen, Axel Legay, Marius Mikucionis, Danny Bøgsted Poulsen, and Sean Sedwards. Statistical model checking for biological systems. *International Journal on Software Tools for Technology Transfer*, 17(3):351, 2015.

[143] Xiangfang Li, Oluwaseyi Omotere, Lijun Qian, and Edward R. Dougherty. Review of stochastic hybrid systems with applications in biological systems modeling and analysis. *EURASIP Journal on Bioinformatics and Systems Biology*, 2017(1):8, Jun 2017.

[144] Joao P Hespanha. Modelling and analysis of stochastic hybrid systems. *IEE Proceedings-Control Theory and Applications*, 153(5):520–535, 2006.

[145] Andrew R Teel, Anantharaman Subbaraman, and Antonino Sferlazza. Stability analysis for stochastic hybrid systems: A survey. *Automatica*, 50(10):2435–2456, 2014.

[146] Jianghai Hu, John Lygeros, and Shankar Sastry. *Towards a Theory of Stochastic Hybrid Systems*, pages 160–173. Springer Berlin Heidelberg, Berlin, Heidelberg, 2000.

[147] Mohammad Soltani and Abhyudai Singh. Stochastic analysis of linear time-invariant systems with renewal transitions. In *American Control Conference (ACC), 2017*, pages 1734–1739. IEEE, 2017.

[148] Mohammad Soltani and Abhyudai Singh. Moment-based analysis of stochastic hybrid systems with renewal transitions. *Automatica*, 84:62–69, 2017.

[149] Leslaw Socha. *Linearization methods for stochastic dynamic systems*, volume 730. Springer Science & Business Media, 2007.

[150] Alessandro Borri, Francesco Carravetta, and Pasquale Palumbo. Cubification of nonlinear stochastic differential equations and approximate moments calculation of the langevin equation. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 4540–4545. IEEE, 2016.

[151] Christian Kuehn. *Moment Closure–A Brief Review*, pages 253–271. Springer International Publishing, Cham, 2016.

[152] Abhyudai Singh and Joao P Hespanha. Approximate moment dynamics for chemically reacting systems. *IEEE Transactions on Automatic Control*, 56(2):414–418, 2011.

[153] Mohammad Soltani, Cesar Augusto Vargas-Garcia, and Abhyudai Singh. Conditional moment closure schemes for studying stochastic dynamics of genetic circuits. *IEEE transactions on biomedical circuits and systems*, 9(4):518–526, 2015.

[154] Ping Wang, Lydia Robert, James Pelletier, Wei Lien Dang, Francois Taddei, Andrew Wright, and Suckjoon Jun. Robust growth of escherichia coli. *Current Biology*, 20:1099–1103, 2010.

[155] Jonathan J. Turner, Jennifer C. Ewald, and Jan M. Skotheim. Cell size control in yeast. *Current Biology*, 22:R350–R359, 2012.

[156] Wallace F. Marshall, Kevin D. Young, Matthew Swaffer, Elizabeth Wood, Paul Nurse, Akatsuki Kimura, Joseph Frankel, John Wallingford, Virginia Walbot, Xian Qu, and Adrienne HK Roeder. What determines cell size? *BMC Biology*, 10:101, 2012.

[157] Lydia Robert, Marc Hoffmann, Nathalie Krell, Stéphane Aymerich, Jérôme Robert, and Marie Doumic. Division in *Escherichia coli* is triggered by a size-sensing rather than a timing mechanism. *BMC Biology*, 12:17, 2014.

[158] Saurabh Modi, Cesar Augusto Vargas-Garcia, Khem Raj Ghusinga, and Abhyudai Singh. Analysis of noise mechanisms in cell-size control. *Biophysical Journal*, 112(11):2408–2418, 2017.

[159] Cesar Augusto Vargas-Garcia, Mohammad Soltani, and Abhyudai Singh. Conditions for cell size homeostasis: A stochastic hybrid systems approach. *IEEE Life Sciences Letters*, 2:47–50, 2016.

[160] William Feller. *An introduction to probability theory and its applications: volume II*. John Wiley & Sons New York, 1968.

[161] Nikolai G Ushakov. *Selected topics in characteristic functions*. Walter de Gruyter, 1999.

[162] Torben Maack Bisgaard and Zoltán Sasvári. *Characteristic Functions and Moment Sequences: Positive Definiteness in Probability*. Nova Publishers, 2000.

[163] Zoltán Sasvári. *Multivariate characteristic and correlation functions*. Walter de Gruyter, 2013.

[164] Mircea Grigoriu. *Stochastic calculus: applications in science and engineering*. Springer Science & Business Media, 2013.

[165] Peter W Glynn, Assaf Zeevi, et al. Bounding stationary expectations of markov processes. In *Markov processes and related topics: a Festschrift for Thomas G. Kurtz*, pages 195–214. Institute of Mathematical Statistics, 2008.

[166] Juan Kuntz, Michela Ottobre, Guy-Bart Stan, and Mauricio Barahona. Bounding stationary averages of polynomial diffusions via semidefinite programming. *SIAM Journal on Scientific Computing*, 38(6):A3891–A3920, 2016.

[167] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[168] Edward Allen. *Modeling with Itô stochastic differential equations*, volume 22. Springer Science & Business Media, 2007.

[169] John J. Dennehy and Ing-Nang Wang. Factors influencing lysis time stochasticity in bacteriophage λ. *BMC Microbiology*, 11:174, 2011.

[170] Rebecca White, Shinobu Chiba, Ting Pang, Jill S. Dewey, Christos G. Savva, Andreas Holzenburg, Kit Pogliano, and Ry Young. Holin triggering in real time. *Proceedings of the National Academy of Sciences*, 108:798–803, 2011.

[171] A. Singh and John J. Dennehy. Stochastic holin expression can account for lysis time variation in the bacteriophage λ. *Journal of the Royal Society Interface*, 11:20140140, 2014.

[172] Long Cai and Nir Friedmanand X. Sunney Xie. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440:358–362, 2006.

[173] Daniel K. Wells, William L. Kath, and Adilson E. Motter. Control of stochastic and induced switching in biophysical networks. *Physical Review X*, 5:031036, 2015.

[174] Adam M Corrigan, Edward Tunnacliffe, Danielle Cannon, and Jonathan R Chubb. A continuum model of transcriptional bursting. *eLife*, 5:e13051, 2016.

[175] R. D. Dar, B. S. Razooky, A. Singh, T. V. Trimeloni, J. M. McCollum, C. D. Cox, M. L. Simpson, and L. S. Weinberger. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences*, 109:17454–17459, 2012.

[176] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef. Mammalian genes are transcribed with widely different bursting kinetics. *Science*, 332:472–474, 2011.

[177] Michael Hinczewski and D. Thirumalai. Cellular signaling networks function as generalized wiener-kolmogorov filters to suppress noise. *Physical Review X*, 4:041017, 2014.

[178] A. Singh, Brandon Razooky, Chris D. Cox, Michael L. Simpson, and Leor S. Weinberger. Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophysical Journal*, 98:L32–L34, 2010.

[179] Evgeny Yurkovsky and Iftach Nachman. Event timing at the single–cell level. *Briefings in Functional Genomics*, 12:90–98, 2013.

[180] J. Paulsson. Model of stochastic gene expression. *Physics of Life Reviews*, 2:157–175, 2005.

[181] Vlad Elgart, Tao Jia, Andrew T Fenley, and Rahul Kulkarni. Connecting protein and mrna burst distributions for stochastic models of gene expression. *Physical Biology*, 8:046001, 2011.

[182] V. Shahrezaei and P. S. Swain. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*, 105:17256–17261, 2008.

[183] Andrea Giometto, Florian Altermatt, Francesco Carrara, Amos Maritan, and Andrea Rinaldo. Scaling body size fluctuations. *Proceedings of the National Academy of Sciences*, 110:4646–4650, 2013.

[184] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. Chapman and Hall/CRC, 2011.

[185] A. Singh and J. P. Hespanha. Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophysical Journal*, 96:4013–4023, 2009.

[186] Arthur T Winfree. *The geometry of biological time*, volume 12. Springer Science & Business Media, 2001.

[187] Vahid Shahrezaei, Julien F. Ollivier, and Peter S. Swain. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology*, 4, 2008.

[188] Abhyudai Singh. Transient changes in intercellular protein variability identify sources of noise in gene expression. *Biophysical Journal*, 107:2214–2220, 2014.

[189] Ing-Nang Wang. Lysis timing and bacteriophage fitness. *Genetics*, 172:17–26, 2006.

[190] Richard H Heineman and James J Bull. Testing optimality with experimental evolution: lysis time in a bacteriophage. *Evolution*, 61:1695–1709, 2007.

[191] Yongping Shao and Ing Nang Wang. Bacteriophage adsorption rate and optimal lysis time. *Genetics*, 180:471–482, 2008.

[192] Juan A Bonachela and Simon A Levin. Evolutionary comparison between viral lysis rate and latent period. *Journal of Theoretical Biology*, 345:32–42, 2014.

[193] Chung-Yu Chang, K Nam, and Ry Young. *S* gene expression and the timing of lysis by bacteriophage lambda. *Journal of Bacteriology*, 177:3283–3294, 1995.

[194] Y. P. Shao and Ing-Nang Wang. Effect of late promoter activity on bacteriophage lambda fitness. *Genetics*, 181:1467–1475, 2009.

[195] Angelika Gründling, David L. Smith, Udo Bläsi, and Ry Young. Dimerization between the holin and holin inhibitor of phage λ. *Journal of Bacteriology*, 182:6075–6081, 2000.

[196] U. Blasi and R. Young. Two beginnings for a single purpose: The dual-start holins in the regulation of phage lysis. *Molecular Microbiolog*, 21:675–682, 1996.

[197] W. J. Voorn and L. J. H. Koppes. Skew or third moment of bacterial generation times. *Archives of Microbiology*, 169:43–51, 1997.

[198] Manuel Campos, Ivan V. Surovtsev, Setsu Kato, Ahmad Paintdakhi, Bruno Beltran, Sarah E. Ebmeier, and Christine Jacobs-Wagner. A constant size extension drives bacterial cell size homeostasis. *Cell*, 159:1433–1446, 2014.

[199] Maxime Deforet, Dave van Ditmarsch, and João B. Xavier. Cell-size homeostasis and the incremental rule in a bacterial pathogen. *Biophysical Journal*, 109:521–528, 2015.

[200] Anouchka Fievet, Adrien Ducret, Tâm Mignot, Odile Valette, Lydia Robert, Romain Pardoux, Alain Roger Dolla, and Corinne Aubert. Single-cell analysis of growth and cell division of the anaerobe Desulfovibrio vulgaris Hildenborough. *Frontiers in Microbiology*, 6:1378, 2015.

[201] RM Teather, JF Collins, and WD Donachie. Quantal behavior of a diffusible factor which initiates septum formation at potential division sites in Escherichia coli. *Journal of Bacteriology*, 118:407–413, 1974.

[202] PA Fantes, WD Grant, RH Pritchard, PE Sudbery, and AE Wheals. The regulation of cell size and the control of mitosis. *Journal of Theoretical Biology*, 50:213–244, 1975.

[203] Markus Basan, Manlu Zhu, Xiongfeng Dai, Mya Warren, Daniel Sévin, Yi-Ping Wang, and Terence Hwa. Inflating bacterial cells by increased protein synthesis. *Molecular Systems Biology*, 11:836, 2015.

[204] W. D. Donachie. Relationship between cell size and time of initiation of DNA replication. *Nature*, 219:1077–1079, 1968.

[205] L. Sompayrac and O. Maaløe. Autorepressor model for control of DNA replication. *Nature*, 241:133–135, 1973.

[206] Po-Yi Ho and Ariel Amir. Simultaneous regulation of cell size and chromosome replication in bacteria. *Microbial Physiology and Metabolism*, 6:662, 2015.

[207] W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422:633–637, 2003.

[208] Arjun Raj and Alexander van Oudenaarden. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135:216–226, 2008.

[209] Abhyudai Singh and Mohammad Soltani. Quantifying intrinsic and extrinsic variability in stochastic gene expression models. *PLOS ONE*, 8:e84301, 2013.

[210] Srividya Iyer-Biswas, Charles S. Wright, Jonathan T. Henry, Klevin Lo, Stanislav Burov, Yihan Lin, Gavin E. Crooks, Sean Crosson, Aaron R. Dinner, and Norbert F. Scherer. Scaling laws governing stochastic growth and division of single bacterial cells. *Proceedings of the National Academy of Sciences*, 111:15912–15917, 2014.

[211] N. Friedman, L. Cai, and X.S. Xie. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Physical Review Letters*, 97:168302, 2006.

[212] Otto G. Berg. A model for the statistical fluctuations of protein numbers in a microbial population. *Journal of Theoretical Biology*, 71:587–603, 1978.

[213] D. R. Rigney. Stochastic model of constitutive protein levels in growing and dividing bacterial cells. *Journal of Theoretical Biology*, 76:453–480, 1979.

[214] Olivia Padovan-Merhar, Gautham P. Nair, Andrew G. Biaesch, Andreas Mayer, Steven Scarfone, Shawn W. Foley, Angela R. Wu, L. Stirling Churchman, Abhyudai Singh, and Arjun Raj. Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Molecular Cell*, 58:339–352, 2015.

[215] Abhyudai Singh and Pavol Bokes. Consequences of mRNA transport on stochastic variability in protein levels. *Biophysical Journal*, 103:1087–1096, 2012.

[216] A. Singh, B. S. Razooky, R. D. Dar, and L. S. Weinberger. Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. *Molecular Systems Biology*, 8:607, 2012.

[217] Stefan Klumpp, Zhongge Zhang, and Terence Hwa. Growth rate-dependent global effects on gene expression in bacteria. *Cell*, 139:1366–1375, 2009.

[218] Marius Hintsche and Stefan Klumpp. Dilution and the theoretical description of growth-rate dependent gene expression. *Journal of Biological Engineering*, 7:22, 2013.

[219] Laurence A. Baxter. Reliability applications of the relevation transform. *Naval Research Logistics*, 29:323–330, 1982.

[220] Franco Pellerey, Moshe Shaked, and Joel Zinn. Nonhomogeneous poisson processes and logconcavity. *Probability in the Engineering and Informational Sciences*, 14:353–373, 2000.

[221] Stephen Cooper and Charles E Helmstetter. Chromosome replication and the division cycle of Escherichia coli B/r. *Journal of Molecular Biology*, 31:519–540, 1968.

[222] Stephen Cooper. *Bacterial growth and division: biochemistry and regulation of prokaryotic and eukaryotic division cycles*. Elsevier, 2012.

[223] Sattar Taheri-Araghi. Self-consistent examination of Donachie's constant initiation size at the single-cell level. *Frontiers in Microbiology*, 6:1349, 2015.

[224] Andrew S Kennard, Matteo Osella, Avelino Javer, Jacopo Grilli, Philippe Nghe, Sander J Tans, Pietro Cicuta, and Marco Cosentino Lagomarsino. Individuality and universality in the growth-division laws of single E. coli cells. *Physical Review E*, 93:012408, 2016.

[225] Srividya Iyer-Biswas, Gavin E Crooks, Norbert F Scherer, and Aaron R Dinner. Universality in stochastic exponential growth. *Physical Review Letters*, 113:028101, 2014.

[226] Matteo Osella, Eileen Nugent, and Marco Cosentino Lagomarsino. Concerted control of *Escherichia coli* cell division. *Proceedings of the National Academy of Sciences*, 111:3431–3435, 2014.

[227] Joe Lutkenhaus. Assembly dynamics of the bacterial MinCDE system and spatial regulation of the Z ring. *Annual Reviews of Biochemisty*, 76:539–562, 2007.

[228] Harold P Erickson, David E Anderson, and Masaki Osawa. FtsZ in bacterial cytokinesis: cytoskeleton and force generator all in one. *Microbiology and Molecular Biology Reviews*, 74:504–528, 2010.

[229] Piet AJ De Boer. Advances in understanding E. coli cell fission. *Current Opinion in Microbiology*, 13:730–737, 2010.

[230] Lydia Robert. Size sensors in bacteria, cell cycle control, and size control. *Frontiers in Microbiology*, 6:515, 2015.

[231] Erfei Bi and J Lutkenhaus. FtsZ regulates frequency of cell division in Escherichia coli. *Journal of Bacteriology*, 172:2765–2768, 1990.

[232] David W Adams and Jeff Errington. Bacterial cell division: assembly, maintenance and disassembly of the Z ring. *Nature Reviews Microbiology*, 7:642–653, 2009.

[233] An-Chun Chien, Norbert S Hill, and Petra Anne Levin. Cell size control in bacteria. *Current Biology*, 22:R340–R349, 2012.

[234] Tove Atlung, Anders Løbner-Olesen, and Flemming G Hansen. Overproduction of dnaa protein stimulates initiation of chromosome and minichromosome replication in Escherichia coli. *Molecular and General Genetics MGG*, 206:51–59, 1987.

[235] Anders Løbner-Olesen, Kirsten Skarstad, Flemming G Hansen, Kaspar von Meyenburg, and Erik Boye. The DnaA protein determines the initiation mass of Escherichia coli K-12. *Cell*, 57:881–889, 1989.

[236] William D Donachie and Garry W Blakely. Coupling the initiation of chromosome replication to cell size in Escherichia coli. *Current Opinion in Microbiology*, 6:146–150, 2003.

[237] Suckjoon Jun and Sattar Taheri-Araghi. Cell-size maintenance: universal strategy revealed. *Trends in Microbiology*, 23:4–6, 2015.

[238] Yu Tanouchi, Anand Pai, Heungwon Park, Shuqiang Huang, Rumen Stamatov, Nicolas E. Buchler, and Lingchong You. A noisy linear map underlies oscillations in cell size and gene expression in bacteria. *Nature*, 523:357–360, 2015.

[239] Xiaoxin Liao, LQ Wang, and Pei Yu. *Stability of Dynamical Systems*, volume 5. Elsevier, 2007.

Appendix A

SEMIDEFINITENESS OF OUTER PRODUCTS AND MOMENT MATRICES

A symmetric $n \times n$ matrix H is said to be positive semidefinite if

$$z^\top H z \geq 0 \tag{A.1}$$

for any $n \times 1$ vector z. Let $H = \mathbf{\Gamma}\mathbf{\Gamma}^\top$ for a $n \times 1$ vector $\mathbf{\Gamma}$ that consists of monomials of $\boldsymbol{x}$ up to a certain order. We have

$$z^\top \mathbf{\Gamma}\mathbf{\Gamma}^\top z = \left(z^\top \mathbf{\Gamma}\right)\left(z^\top \mathbf{\Gamma}\right)^\top = \left\|z^\top \mathbf{\Gamma}\right\|^2 \geq 0. \tag{A.2}$$

Therefore, the outer product $\mathbf{\Gamma}\mathbf{\Gamma}^\top$ is always positive semidefinite. Recall that the moment matrix H_e is generated by taking expectation of the outer product $\mathbf{\Gamma}\mathbf{\Gamma}^\top$. Its positive semidefiniteness can be proved as follows

$$z^\top \left\langle \mathbf{\Gamma}\mathbf{\Gamma}^\top \right\rangle z = \left\langle \left(z^\top \mathbf{\Gamma}\right)\left(z^\top \mathbf{\Gamma}\right)^\top \right\rangle = \left\langle \left\|z^\top \mathbf{\Gamma}\right\|^2 \right\rangle \geq 0. \tag{A.3}$$

Same arguments can be used to prove semidefiniteness of the outer product $h_i(\boldsymbol{x})\mathbf{\Gamma}\mathbf{\Gamma}^\top$ and the moment matrix H_{o_i} obtained by taking expectation thereof.

Appendix B

SUPPLEMENTARY DETAILS FOR CHAPTER 4

B.1 Some Properties of Matrix Λ

The matrix Λ is given by

$$\Lambda = \begin{bmatrix} -\frac{bk_0}{b+1} & \gamma & \cdots & 0 \\ \frac{bk_0}{(b+1)^2} & -\left(\frac{bk_1}{b+1}+\gamma\right) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{b^{\theta-2}k_0}{(b+1)^{\theta-1}} & \frac{b^{\theta-3}k_1}{(b+1)^{\theta-2}} & \cdots & (\theta-1)\gamma \\ \frac{b^{\theta-1}k_0}{(b+1)^{\theta}} & \frac{b^{\theta-2}k_1}{(b+1)^{\theta-1}} & \cdots & -\left(\frac{bk_{\theta-1}}{b+1}+(\theta-1)\gamma\right) \end{bmatrix}. \tag{B.1}$$

In order to prove that Λ is a Hurwitz matrix, we show that the following two conditions hold true [239, pp. 48–49]:

1. The diagonal elements $a_{ii} < 0$ for $i = 1, 2, \cdots, \theta$,
2. $\max\limits_{1\le j\le\theta} \sum\limits_{\substack{i=1 \\ j\neq i}}^{\theta} \left|\frac{a_{ij}}{a_{jj}}\right| < 1.$

The first requirement above is clearly satisfied as $a_{ii} = -\frac{bk_{i-1}}{b+1} - (i-1)\gamma < 0$. To check whether the second requirement is satisfied, for any column $j = 1, 2, \cdots, \theta$ we have that

$$\sum_{\substack{i=1 \\ j\neq i}}^{\theta} \left|\frac{a_{ij}}{a_{jj}}\right| = \frac{(j-1)\gamma}{\frac{bk_{j-1}}{b+1}+(j-1)\gamma} + \frac{bk_{j-1}}{b+1} \sum_{i=j+1}^{\theta} \frac{\left(\frac{b}{b+1}\right)^{i-j}}{\frac{bk_{j-1}}{b+1}+(j-1)\gamma} \tag{B.2a}$$

$$= \frac{(j-1)\gamma}{\frac{bk_{j-1}}{b+1}+(j-1)\gamma} + \frac{\frac{bk_{j-1}}{b+1}\left(1-\left(\frac{b}{b+1}\right)^{\theta-j}\right)}{\frac{bk_{j-1}}{b+1}+(j-1)\gamma} \tag{B.2b}$$

$$= \frac{\frac{bk_{j-1}}{b+1}\left(1-\left(\frac{b}{b+1}\right)^{\theta-j}\right)+(j-1)\gamma}{\frac{bk_{j-1}}{b+1}+(j-1)\gamma} < 1. \tag{B.2c}$$

Thus, the matrix Λ is Hurwitz, i.e., the eigenvalues of Λ have negative real part.

B.2 Optimal Feedback when Protein Does Not Degrade

As mentioned in the main text, our objective is to find optimal feedback strategy that minimizes $\langle \mathbf{T}^2 \rangle$ such that $\langle \mathbf{T} \rangle$ is fixed. For calculation purposes, we will denote this constraint as $\langle \mathbf{T} \rangle = t_{opt}$. Let ϵ represents the Lagrange's multiplier, then we define the following objective function

$$\Pi := \left\langle \mathbf{T}^2 \right\rangle + \epsilon \left(\langle \mathbf{T} \rangle - t_{opt} \right). \tag{B.3}$$

The optimization problem is solved in two steps. First, we determine the critical points. Second, we find the critical point corresponding to a global minimum.

Determining the critical points requires the following system of equations to be solved

$$\frac{\partial \langle \mathbf{T}^2 \rangle}{\partial k_i} = \epsilon \frac{\partial \langle \mathbf{T} \rangle}{\partial k_i}, 0 \leq i \leq \theta - 1, \tag{B.4a}$$

$$\langle \mathbf{T} \rangle = t_{opt}. \tag{B.4b}$$

The expressions of the first-two moments of FPT when protein does not degrade can be obtained by substituting $\gamma = 0$ in the moment expressions. These are given by

$$\langle \mathbf{T} \rangle = \frac{1}{k_0} + \frac{1}{b} \sum_{i=0}^{\theta-1} \frac{1}{k_i}, \tag{B.5a}$$

$$\left\langle \mathbf{T}^2 \right\rangle = \frac{2}{b^2} \left(\frac{\tau_0}{b k_0} + \sum_{i=0}^{\theta-1} \frac{\tau_i}{k_i} \right), \text{ where } \tau_i := \frac{b}{k_i} + \sum_{j=i}^{\theta-1} \frac{1}{k_j}. \tag{B.5b}$$

The optimization problem in (B.4) requires calculation of the first order derivatives of $\langle \mathbf{T} \rangle$ and $\langle \mathbf{T}^2 \rangle$. The derivatives of $\langle \mathbf{T} \rangle$ with respect to k_i's are given by

$$\frac{\partial \langle \mathbf{T} \rangle}{\partial k_0} = \frac{b+1}{b} \left(-\frac{1}{k_0^2} \right); \frac{\partial \langle \mathbf{T} \rangle}{\partial k_i} = \frac{1}{b} \left(-\frac{1}{k_i^2} \right), \ 1 \leq i \leq \theta - 1. \tag{B.6}$$

Similarly, the derivative of $\langle \mathbf{T}^2 \rangle$ are

$$\frac{\partial \langle \mathbf{T}^2 \rangle}{\partial k_0} = 2 \left(\frac{1}{b} \right)^2 \left(-\frac{2(b+1)^2}{k_0^3} - \frac{b+1}{k_0^2} \sum_{j=1}^{\theta-1} \frac{1}{k_j} \right), \tag{B.7a}$$

$$\frac{\partial \langle \mathbf{T}^2 \rangle}{\partial k_i} = 2 \left(\frac{1}{b} \right)^2 \left(\frac{b+1}{k_0} + \frac{2(b+1)}{k_i} + \sum_{\substack{j=1 \\ j \neq i}}^{\theta-1} \frac{1}{k_j} \right) \left(\frac{-1}{k_i^2} \right), 1 \leq i \leq \theta - 1. \tag{B.7b}$$

Substituting these expressions and assuming $k_0 \neq 0$, $k_i \neq 0$, the system of equations to be solved becomes

$$\frac{2}{b}\left(\frac{2\,(b+1)}{k_0}+\sum_{j=1}^{\theta-1}\frac{1}{k_j}\right)=\epsilon, \tag{B.8a}$$

$$\frac{2}{b}\left(\frac{b+1}{k_0}+\frac{2\,(b+1)}{k_i}+\sum_{\substack{j=1\\ j\neq i}}^{\theta-1}\frac{1}{k_j}\right)=\epsilon, 1\leq i\leq \theta-1, \tag{B.8b}$$

$$\frac{1}{b}\left(\frac{b+1}{k_0}+\sum_{j=1}^{\theta-1}\frac{1}{k_j}\right)=t_{opt}. \tag{B.8c}$$

Solution to these equations is given by

$$k_0=\frac{(b+1)(2b+\theta)}{(2b+1)bt_{opt}} \tag{B.9a}$$

$$k_i=\frac{2b+1}{b+1}k_0=\frac{2b+\theta}{bt_{opt}}, 1\leq i\leq \theta-1, \tag{B.9b}$$

$$\epsilon=\frac{2\,(b+1)\,(4b+\theta+1)}{(2b+1)\,bk_0}=\frac{2t_{opt}(4b+\theta+1)}{2b+\theta}. \tag{B.9c}$$

We have calculated the critical point for the optimization problem. However, it needs to be checked whether its an minimum or maximum. For this purpose, we consider the bordered Hessian as follows.

$$D_\Pi=\begin{bmatrix} \frac{\partial^2\Pi}{\partial\epsilon^2} & \frac{\partial^2\Pi}{\partial\epsilon\partial k_0} & \frac{\partial^2\Pi}{\partial\epsilon\partial k_1} & \cdots & \frac{\partial^2\Pi}{\partial\epsilon\partial k_{\theta-1}} \\ \frac{\partial^2\Pi}{\partial k_0\partial\epsilon} & \frac{\partial^2\Pi}{\partial k_0^2} & \frac{\partial^2\Pi}{\partial k_0\partial k_1} & \cdots & \frac{\partial^2\Pi}{\partial k_0\partial k_{\theta-1}} \\ \frac{\partial^2\Pi}{\partial k_1\partial\epsilon} & \frac{\partial^2\Pi}{\partial k_1\partial k_0} & \frac{\partial^2\Pi}{\partial k_1^2} & \cdots & \frac{\partial^2\Pi}{\partial k_1\partial k_{\theta-1}} \\ \vdots & \vdots & \vdots & \vdots & \\ \frac{\partial^2\Pi}{\partial k_{\theta-1}\partial\epsilon} & \frac{\partial^2\Pi}{\partial k_{\theta-1}\partial k_0} & \frac{\partial^2\Pi}{\partial k_{\theta-1}\partial k_1} & \cdots & \frac{\partial^2\Pi}{\partial k_{\theta-1}^2} \end{bmatrix}. \tag{B.10}$$

We will show that all the principal minors of this matrix are negative. To start with, let us first determine the second order derivatives of $\langle\mathbf{T}\rangle$.

$$\frac{\partial^2\langle\mathbf{T}\rangle}{\partial k_0^2}=\frac{2\,(b+1)}{bk_0^3};\quad \frac{\partial^2\langle\mathbf{T}\rangle}{\partial k_i^2}=\frac{2}{bk_i^3}, 1\leq i\leq\theta-1;\quad \frac{\partial^2\langle\mathbf{T}\rangle}{\partial k_i\partial k_j}=0, 0\leq i,j\leq\theta-1, i\neq j. \tag{B.11}$$

Similarly, the derivatives for $\langle \mathbf{T}^2 \rangle$ are given by

$$\frac{\partial^2 \langle \mathbf{T}^2 \rangle}{\partial k_0^2} = 2\left(\frac{1}{b}\right)^2 \left(\frac{6\,(b+1)^2}{k_0^4} + \frac{2\,(b+1)}{k_0^3}\sum_{j=1}^{\theta-1}\frac{1}{k_j}\right), \tag{B.12a}$$

$$\frac{\partial^2 \langle \mathbf{T}^2 \rangle}{\partial k_i^2} = 2\left(\frac{1}{b}\right)^2 \left(\frac{2\,(b+1)}{k_0 k_i^3} + \frac{6\,(b+1)}{k_i^4} + \frac{2}{k_i^3}\sum_{j=1}^{\theta-1}\frac{1}{k_j}\right), 1 \le i \le \theta-1, \tag{B.12b}$$

$$\frac{\partial^2 \langle \mathbf{T}^2 \rangle}{\partial k_0 \partial k_i} = \frac{\partial^2 \langle \mathbf{T}^2 \rangle}{\partial k_i \partial k_0} = 2\left(\frac{1}{b}\right)^2 \frac{b+1}{k_0^2 k_i^2}, 1 \le i \le \theta-1, \tag{B.12c}$$

$$\frac{\partial^2 \langle \mathbf{T}^2 \rangle}{\partial k_i \partial k_j} = 2\left(\frac{1}{b}\right)^2 \frac{1}{k_i^2 k_j^2}, 1 \le i, j \le \theta-1, i \ne j. \tag{B.12d}$$

We can now determine the elements of the bordered Hessian matrix in equation (B.10) computed at the solution given by equations (B.9)

$$\frac{\partial^2 \Pi}{\partial \epsilon^2} = \frac{\partial^2 \left(\langle T^2 \rangle + \epsilon\left(\langle T \rangle - t_{opt}\right)\right)}{\partial \epsilon^2} = 0, \tag{B.13a}$$

$$\frac{\partial^2 \Pi}{\partial \epsilon \partial k_0} = \frac{\partial^2 \Pi}{\partial k_0 \partial \epsilon} = \frac{\partial \langle T \rangle}{\partial k_0} = \frac{b+1}{b}\left(-\frac{1}{k_0^2}\right). \tag{B.13b}$$

$$\frac{\partial^2 \Pi}{\partial \epsilon \partial k_i} = \frac{\partial^2 \Pi}{\partial k_i \partial \epsilon} = \frac{\partial \langle T \rangle}{\partial k_i} = \frac{1}{b}\left(-\frac{1}{k_i^2}\right). \tag{B.13c}$$

$$\frac{\partial^2 \Pi}{\partial k_0^2} = \frac{\partial^2 \left(\langle T^2 \rangle + \epsilon\left(\langle T \rangle - t_{opt}\right)\right)}{\partial k_0^2} = \frac{4(b+1)^2(10b+2\theta+3)}{(2b+1)b^2 {k_0}^4} \tag{B.13d}$$

$$\frac{\partial^2 \Pi}{\partial k_i^2} = \frac{\partial^2 \left(\langle T^2 \rangle + \epsilon\left(\langle T \rangle - t_{opt}\right)\right)}{\partial k_i^2} = \frac{4(b+1)^4(9b+2\theta+4)}{(2b+1)^4 b^2 {k_0}^4}. \tag{B.13e}$$

$$\frac{\partial^2 \Pi}{\partial k_0 k_i} = \frac{\partial^2 \Pi}{\partial k_i k_0} = \frac{\partial^2 \left(\langle T^2 \rangle + \epsilon\left(\langle T \rangle - t_{opt}\right)\right)}{\partial k_0 \partial k_i} = 2\left(\frac{1}{b}\right)^2 \frac{b+1}{k_0^2 k_i^2} \tag{B.13f}$$

$$\frac{\partial^2 \Pi}{\partial k_j k_i} = \frac{\partial^2 \Pi}{\partial k_i k_j} = \frac{\partial^2 \left(\langle T^2 \rangle + \epsilon\left(\langle T \rangle - t_{opt}\right)\right)}{\partial k_j \partial k_i} = 2\left(\frac{1}{b}\right)^2 \frac{1}{k_j^2 k_i^2} \tag{B.13g}$$

It can be noted elements of D_Π are from a set six quantities. Defining

$$q_1 := \frac{b+1}{b}\left(-\frac{1}{k_0^2}\right), \qquad q_2 := \frac{1}{b}\left(-\frac{1}{k_0^2}\right), \tag{B.14a}$$

$$q_3 := \frac{4(b+1)^2(10b+2\theta+3)}{(2b+1)b^2{k_0}^4}, \quad q_4 := \frac{2(b+1)^3}{(2b+1)^2b^2{k_0}^4}, \tag{B.14b}$$

$$q_5 := \frac{4(b+1)^4(9b+2\theta+4)}{(2b+1)^4b^2{k_0}^4}, \quad q_6 := \frac{2\left(b+1\right)^2}{\left(2b+1\right)^2b^2k_0^4} \tag{B.14c}$$

we can write D_Π as

$$D_\Pi = \begin{bmatrix} 0 & q_1 & q_2 & \cdots & q_2 \\ q_1 & q_3 & q_4 & \cdots & q_4 \\ q_2 & q_4 & q_5 & \cdots & q_6 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_2 & q_4 & q_6 & \cdots & q_5 \end{bmatrix}. \tag{B.15}$$

Let us denote by $\mathcal{M}(n)$ the principal minor of the matrix D_Π of size $n \times n$. It can be easily seen that $\mathcal{M}(1) = 0$, $\mathcal{M}(2) = 0 - q_1^2 < 0$ and $\mathcal{M}(3) = -q_2^2q_3 + 2q_1q_2q_4 - q_1^2q_5$. For $4 \leq n \leq \theta$, we perform the following two elementary operations on D_Π

- $col_r = col_r - col_{r-1}$
- $row_r = row_r - row_{r-1}$

for $r = n, n-1, ..., 1$. This yields

$$\mathcal{M}(n) = 2(q_5 - q_6)\mathcal{M}(n-1) - (q_5 - q_6)^2\mathcal{M}(n-2), \quad 4 \leq n \leq \theta. \tag{B.16}$$

The solution to the above recursive equation is given by

$$\mathcal{M}(n) = -(q_5 - q_6)^{n-3}\left((n-2)\left(q_2^2q_3 - 2q_1q_2q_4 + q_1^2q_6\right) + \left(q_1^2q_6 + q_1^2q_5\right)\right). \tag{B.17}$$

It can be easily checked that $\mathcal{M}(n)$ is negative because $q_5 > q_6$, $q_2^2q_3 - 2q_1q_2q_4 + q_1^2q_6 > 0$ and $q_1^2q_6 + q_1^2q_5 > 0$. This proves that the critical point indeed corresponds to a minimum.

B.3 Optimal Feedback when Protein Does Not Degrade in Presence of Extrinsic Noise

In this section, assuming that the protein does not degrade, we investigate how the optimal regulation strategy deviates from a no feedback in presence of a static

extrinsic noise. We consider that the extrinsic noise affects the transcription rate. To this end, we assume that a factor $\boldsymbol{\zeta}$ multiplies with the transcription rates resulting in an effective transcription rate when $\boldsymbol{x}(t) = i$ to be $k_i\boldsymbol{\zeta}$.

We consider an extrinsic factor $\boldsymbol{\zeta}$ with a positive-valued arbitrary distribution $\rho_{\boldsymbol{\zeta}}(\zeta)$. This factor is assumed to be static, i.e., it does not vary over the time scale of the event of interest. Further we assume that it affects the transcription rates in a multiplicative fashion. The first-passage time mean in this case can be written as

$$\langle \mathbf{T}|\boldsymbol{\zeta} = \zeta\rangle = \frac{1}{\zeta}\left(\frac{1}{k_0} + \frac{1}{b}\sum_{i=0}^{\theta-1}\frac{1}{k_i}\right) \implies \langle \mathbf{T}\rangle = \left\langle\frac{1}{\boldsymbol{\zeta}}\right\rangle\left(\frac{1}{k_0} + \frac{1}{b}\sum_{i=0}^{\theta-1}\frac{1}{k_i}\right). \tag{B.18}$$

Likewise the second order moment can be written as

$$\left\langle \mathbf{T}^2\right\rangle = \left\langle\frac{1}{\boldsymbol{\zeta}^2}\right\rangle\frac{2}{b^2}\left(\frac{\tau_0}{bk_0} + \sum_{i=0}^{\theta-1}\frac{\tau_i}{k_i}\right), \tau_i := \frac{b}{k_i} + \sum_{j=i}^{\theta-1}\frac{1}{k_j}. \tag{B.19}$$

Solving the constrained optimization problem of minimizing $\langle \mathbf{T}^2\rangle$ constrained to $\langle \mathbf{T}\rangle = t_{opt}$ in this case simplifies to solving the following system of equations

$$\frac{\left\langle 1/\boldsymbol{\zeta}^2\right\rangle}{\langle 1/\boldsymbol{\zeta}\rangle}\frac{2}{b}\left(\frac{2\,(b+1)}{k_0} + \sum_{j=1}^{\theta-1}\frac{1}{k_j}\right) = \epsilon, \tag{B.20a}$$

$$\frac{\left\langle 1/\boldsymbol{\zeta}^2\right\rangle}{\langle 1/\boldsymbol{\zeta}\rangle}\frac{2}{b}\left(\frac{b+1}{k_0} + \frac{2\,(b+1)}{k_i} + \sum_{\substack{j=1\\ j\neq i}}^{\theta-1}\frac{1}{k_j}\right) = \epsilon, 1 \leq i \leq \theta-1, \tag{B.20b}$$

$$\left\langle\frac{1}{\boldsymbol{\zeta}}\right\rangle\frac{1}{b}\left(\frac{b+1}{k_0} + \sum_{j=1}^{\theta-1}\frac{1}{k_j}\right) = t_{opt}, \tag{B.20c}$$

where ϵ represents the Lagrange's multiplier. Solution to these equations gives

$$k_0 = \frac{(b+1)(2b+\theta)\left\langle\frac{1}{\boldsymbol{\zeta}}\right\rangle}{(2b+1)bt_{opt}} \tag{B.21a}$$

$$k_i = \frac{2b+1}{b+1}k_0 = \frac{(2b+\theta)\left\langle\frac{1}{\boldsymbol{\zeta}}\right\rangle}{bt_{opt}}, 1 \leq i \leq \theta-1. \tag{B.21b}$$

Appendix C

DISCLAIMER

This thesis is based on the material that has either been published in peer-reviewed publications [90,91,93–101] or is undergoing peer-review [92]. Below I provide the list of papers for each chapter.

- Chapter 2: [90, 91]
- Chapter 3: [91–94]
- Chapter 4: [95–100]
- Chapter 5: [101]