

UNIVERSITY of CALIFORNIA

Santa Barbara

**Stochastic modeling of chemical reactions and gene regulatory networks**

A Dissertation submitted in partial satisfaction of the  
requirements for the degree

Doctor of Philosophy

in

Electrical and Computer Engineering

by

Abhyudai Singh

Committee in charge:

Professor João Pedro Hespanha, Chair

Professor Mustafa Khammash

Professor Frank Doyle

Professor Roger Nisbet

December 2008

The dissertation of Abhyudai Singh is approved.

---

Professor Mustafa Khammash

---

Professor Frank Doyle

---

Professor Roger Nisbet

---

Professor João Pedro Hespanha, Committee Chair

September 2008

To my parents Yatindra and Neeta Singh

## **Acknowledgements**

I would like to express my deepest gratitude to my advisor João Hespanha for his guidance and insight, without which this thesis would not have been possible.

I am grateful to Roger Nisbet, Mustafa Khammash and Frank Doyle for serving on my Ph.D. committee and for many fruitful discussions on this research.

I would like to thank all my friends in Santa Barbara for making my stay at UCSB a memorable one.

Last but not the least, I thank my grandparents (V. K. S. Chaudhary, G. B. Singh and K. D. Singh), my parents (Yatindra and Neeta Singh) and my lovely wife (Swati Singh). They have all been a great source of strength and inspiration throughout this work.

## Curriculum Vitae

Abhyudai Singh

Abhyudai Singh was born in Allahabad, India, on December 3, 1979.

### Education

2006	M.S. Mechanical Engineering, Michigan State University, East Lansing.
2004	M.S. Electrical and Computer Engineering, Michigan State University, East Lansing.
2002	B.Tech. Mechanical Engineering, Indian Institute of Technology, Kanpur, India.

### Experience

2004–present	Research Assistant, University of California, Santa Barbara.
2003-2004	Research Assistant, Michigan State University, East Lansing.
2002–2003	Teaching Assistant, Michigan State University, East Lansing.

### Selected Publications

**A. Singh**, R. Mukherjee, K. Turner and S. Shaw. *MEMS Implementation of Axial and Follower End Forces*. Journal of Sound and Vibration, 286, 637–644, 2005.

**A. Singh** and H. K. Khalil. *Regulation of Nonlinear Systems Using Conditional Integrators*. International Journal of Robust and Nonlinear Control, 15, 339-362, 2005.

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.

**A. Singh** and J. P. Hespanha. *Lognormal Moment Closures for Chemically Reacting Systems*. In Proc. of the 45th IEEE Conference on Decision and Control, San Diego, 2006.

**A. Singh** and R. M. Nisbet. *Semi-discrete Host-Parasitoid Models*. Journal of Theoretical Biology, 247, 733-742, 2007.

**A. Singh** and J. P. Hespanha. *A Derivative Matching Approach to Moment Closure for the Stochastic Logistic Model*. Bulletin of Math Biology, 69, 1909-1925, 2007.

**A. Singh** and J. P. Hespanha. *Stochastic Analysis of Gene Regulatory Networks Using Moment Closure*. In Proc. of the 2007 American Control Conference, New York, 2007.

## **Abstract**

Stochastic modeling of chemical reactions and gene regulatory networks

by

Abhyudai Singh

Living cells are characterized by small populations of key components (for example, proteins and mRNAs), which make bio-chemical reactions inherently noisy. This thesis outlines new computational techniques for quantifying noise in such bio-chemical reactions. These techniques include a novel moment closure procedure that provides the time evolution of all lower order statistical moments (for example, means and standard deviations) for the number of molecules of different species involved in the reaction. Striking features of this moment closure procedure is that it is independent of the reaction parameters (reaction rates and stoichiometry) and its accuracy can be improved by incurring more computational cost.

This thesis also proposes a new small noise approximation that provides analytical formulas relating the steady-state statistical moments to the parameters of the chemical reaction. Unlike the well-known linear noise approximation, these formulas not only predict the stochastic fluctuations about the mean but also the deviation of the mean from the solution of the corresponding chemical rate equation (i.e., deterministic model).

Using the above computational techniques this thesis investigates the noise suppression properties of various gene network motifs. One such common network motif is

an auto-regulatory negative feedback loop, where the protein expressed from a gene inhibits its own expression. In this network, stochastic fluctuations in protein levels are attributed to two factors: intrinsic noise (i.e., the randomness associated with protein expression and degradation ) and extrinsic noise (i.e., the noise caused by fluctuations in cellular components such as enzyme levels and gene-copy numbers). This thesis shows that although negative feedback loops attenuate both components of noise, they are much more efficient in reducing the extrinsic component of noise than the intrinsic component. It further shows that in these auto-regulatory networks, the protein noise levels are *minimized* at an optimal level of feedback strength. Analytical expressions for this highest level of noise suppression and the amount of feedback that achieves this minimal noise are provided. These theoretical results are shown to be consistent and explain recent experimental observations.

Finally, this thesis examines other common network motifs such as gene cascades, which can act as noise attenuators or noise magnifiers depending upon the amount of intrinsic and extrinsic noise present.



# Contents

<b>Acknowledgements</b>	<b>iv</b>
<b>Curriculum Vitæ</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>List of Figures</b>	<b>xiii</b>
<b>1 Moment closure for stochastic chemical kinetics</b>	<b>1</b>
1.1 Moment dynamics of chemically reacting systems . . . . .	5
1.2 SDM moment closure procedure . . . . .	9
1.2.1 Derivative matching for a generator . . . . .	9
1.2.2 Proof of Theorem 1.2 . . . . .	13
1.2.3 SDM moment closure functions . . . . .	15
1.3 Example . . . . .	19
1.4 Conclusion . . . . .	24
1.4.1 Future work . . . . .	26
<b>2 Small noise approximation</b>	<b>28</b>
2.1 Theoretical formulation . . . . .	29
2.2 Gene expression and protein degradation . . . . .	31
2.2.1 Small noise approximation . . . . .	32
2.2.2 Comparison with linear noise approximation . . . . .	33

2.3	Gene expression and activation . . . . .	34
2.3.1	Small noise approximation . . . . .	37
2.3.2	Comparison with linear noise approximation . . . . .	38
2.4	Conclusion . . . . .	39
<b>3</b>	<b>Moment closure for the stochastic logistic model</b>	<b>40</b>
3.1	Stochastic logistic model . . . . .	43
3.1.1	Model formulation . . . . .	43
3.1.2	Stationary and quasi-stationary distributions . . . . .	44
3.1.3	Transient distributions . . . . .	45
3.2	Time evolution of moments . . . . .	46
3.3	Separable derivative-matching moment closure . . . . .	48
3.4	Distribution based moment closure . . . . .	50
3.4.1	Techniques for obtaining $\phi_{k+1}^{\mathbb{D}}$ . . . . .	51
3.4.2	Cumulant closure functions . . . . .	54
3.5	Comparison of transient performance of moment closures . . . . .	55
3.5.1	Moment closures for $k = 2$ . . . . .	55
3.5.2	Moment closures for $k = 3$ . . . . .	58
3.6	Steady-state solutions of the truncated moment dynamics . . . . .	62
3.7	Conclusion . . . . .	63
3.7.1	Future work . . . . .	64
<b>4</b>	<b>Optimal feedback strength for noise suppression in auto-regulatory gene networks</b>	<b>65</b>
4.1	Un-regulated gene expression . . . . .	71
4.2	Auto-regulatory gene expression . . . . .	73
4.2.1	Linear transcriptional response . . . . .	74
4.2.2	Lambda repressor gene network . . . . .	79
4.2.3	Effect of nonlinearities . . . . .	80
4.3	Extrinsic and intrinsic contributions to noise . . . . .	81
4.4	Auto-regulatory gene networks with negative feedback . . . . .	84

4.4.1	Suppression of intrinsic noise in the protein . . . . .	86
4.4.2	Suppression of extrinsic noise in the protein . . . . .	89
4.4.3	Suppression of total noise in the protein . . . . .	90
4.5	Experimental verification . . . . .	92
4.6	Discussion . . . . .	95
4.6.1	Noise dependence on the shape of the transcriptional response	96
4.6.2	Intrinsic v.s. extrinsic noise . . . . .	96
4.6.3	U-shaped protein noise profile . . . . .	97
4.6.4	Limit of noise suppression . . . . .	98
4.6.5	Positive feedback loops . . . . .	100
4.7	Future work . . . . .	101
<b>5</b>	<b>Evolution of auto-regulatory gene networks</b>	<b>103</b>
5.1	Fitness function . . . . .	105
5.2	Evolvability of the negative feedback . . . . .	106
5.3	Discussion . . . . .	111
5.3.1	Future work . . . . .	113
<b>6</b>	<b>Scaling of stochasticity in gene cascades</b>	<b>115</b>
6.1	Noise in a cascade of genes . . . . .	117
6.2	Scaling of noise in gene cascades . . . . .	119
6.3	Scaling of extrinsic and intrinsic noise in gene cascades . . . . .	122
6.4	Conclusion . . . . .	124
6.4.1	Future work . . . . .	125
	<b>Bibliography</b>	<b>126</b>
<b>A</b>	<b>Moment closure for chemical reactions</b>	<b>134</b>
<b>B</b>	<b>Optimal feedback strength</b>	<b>137</b>
B.1	Effects of nonlinearities in $g(\mathbf{x})$ . . . . .	137
B.2	Extrinsic and intrinsic contributions of noise . . . . .	139

B.3	Limit of noise suppression . . . . .	141
B.4	Estimating the noise in the exogenous signal . . . . .	142
B.5	Gene expression model with transcription and translation . . . . .	143
B.6	Auto-regulatory gene network with transcription and translation . . .	146
B.7	Minimum limit of noise suppression with mRNA dynamics . . . . .	147
<b>C</b>	<b>Incorporating extrinsic noise in the cascade of genes</b>	<b>148</b>

# List of Figures

1.1	Plots of the absolute value of the error between the actual and approximated moment dynamics for a) the mean of $\mathbf{x}_1$ and b) the standard deviation of $\mathbf{x}_1$ corresponding to a second order truncation (solid line) and a third order truncation (dashed line) when the population size is small ( $V = 1$ units). DM and N, refer to the errors corresponding to the SDM moment closure and the normal moment closure, respectively. Other parameters taken as $N_1 = 40$ , $N_2 = 15$ and initial conditions $\mathbf{x}_1(0) = \mathbf{x}_2(0) = 40$ . . . . .	20
1.2	Plots of the absolute value of the error between the actual and approximated moment dynamics for a) the mean of $\mathbf{x}_2$ and b) the standard deviation of $\mathbf{x}_2$ corresponding to a second order truncation (solid line) and a third order truncation (dashed line) when the population size is small ( $V = 1$ units). DM and N, refer to the errors corresponding to the SDM moment closure and the normal moment closure, respectively. Other parameters taken as $N_1 = 40$ , $N_2 = 15$ and initial conditions $\mathbf{x}_1(0) = \mathbf{x}_2(0) = 40$ . . . . .	21
1.3	The steady-states histogram for $\mathbf{x}_1$ , the number of molecules of protein $X_1$ . The solid and dashed line corresponds to a lognormal and normal distribution approximation of the histogram, respectively, with mean and standard deviations obtained from the corresponding truncated moment dynamics . . . . .	22

1.4	Plots of the absolute value of the error between the actual and approximated moment dynamics for a) the mean of $\mathbf{x}_1$ and b) the standard deviation of $\mathbf{x}_1$ corresponding to a second order truncation (solid line) and a third order truncation (dashed line) when the population size is large ( $V = 5$ units). DM and N, refer to the errors corresponding to the SDM moment closure and the normal moment closure, respectively. Other parameters taken as $N_1 = 40$ , $N_2 = 15$ and initial conditions $\mathbf{x}_1(0) = \mathbf{x}_2(0) = 40$ . . . . .	23
2.1	Steady-state histogram of $\mathbf{x}$ obtained from 10,000 Monte Carlo simulations. Solid line is a lognormal approximation using the steady-state moments from the small noise approximation (equation (2.19)). Dashed line is a normal approximation using the steady-state moments from the linear noise approximation (equation (2.22)) . . . . .	36
3.1	Evolutions of the mean error $\mu_1 - v_1$ and of the variance error $(\mu_2 - \mu_1^2) - (v_2 - v_1^2)$ for the different moment closure functions in Table 3.2 for $k = 2$ , with parameters as in (3.51) and $x_0 = 5$ . . . . .	58
3.2	Evolutions of the mean error $\mu_1 - v_1$ and of the variance error $(\mu_2 - \mu_1^2) - (v_2 - v_1^2)$ for the different moment closure functions in Table 3.2 for $k = 2$ , with parameters as in (3.51) and $x_0 = 20$ . . . . .	59
3.3	Evolutions of the mean error $\mu_1 - v_1$ , variance error $(\mu_2 - \mu_1^2) - (v_2 - v_1^2)$ for the different moment closure functions (3.53)–(3.54b) for $k = 3$ , with parameters as in (3.51) and $x_0 = 20$ . . . . .	61
4.1	An auto-regulatory gene network. . . . .	74
4.2	A graphical interpretation of the quantity $I = g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)$ in (4.20) for any arbitrary transcriptional response $g(\mathbf{x})$ : $I$ is the intercept of the tangent to the transcriptional response $g(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}^*$ with the y-axis. . . . .	77
4.3	Time evolution of the average number of protein molecules. 1) corresponds to the case when there is negative feedback (i.e., protein repressed its own transcription) and 2) corresponds to the case when there is gene expression with no negative feedback. The solid and dashed lines represent experimentally measured and fitted approximations to the time evolution of the average number of protein molecules, respectively. This figure was taken from [3]. . . . .	78

4.4	$g_1(\mathbf{x})$ is the standard transcriptional response of a gene network with positive feedback while $g_2(\mathbf{x})$ is the observed transcriptional response in case of the gene in lambda phage encoding the protein lambda repressor. $I_1$ and $I_2$ is the y-intercept of the tangent to the corresponding transcriptional response $g(\mathbf{x})$ at $(\mathbf{x}^*, g(\mathbf{x}^*))$ . . . . .	80
4.5	Intrinsic noise $CV_{int}$ in the protein as a function of the feedback strength $a$ and Hill coefficient $M$ . $CV_{int}$ is normalized by $CV_{int-nr}$ , the intrinsic noise in the protein when there is no feedback. Other parameters taken as $g_0 = 1$ , $b = 0$ , $N_x = 1$ , $V_x = 0$ and $d_x = 0.01$ . . . . .	87
4.6	Total noise $CV_{tot}$ as a function of the feedback strength $a$ when the Hill coefficient is one ( $M = 1$ ) for different values of noise $CV_z$ in the exogenous signal. $CV_{tot}$ is normalized by $CV_{tot-nr}$ , the total noise in the protein when there is no feedback. Other parameters are taken as $g_0 = 1$ , $N_x = 4$ , $V_x^2 = N_x^2 - N_x$ , $b = 0$ , $S = 1$ , $d_x = 0.04$ . The response time $T_z$ is assumed to be much larger than $T_{nr}$ . . . . .	92
4.7	Total noise $CV_{tot}$ as a function of the feedback strength $a$ when the Hill coefficient is two ( $M = 2$ ) for different values of noise $CV_z$ in the exogenous signal. $CV_{tot}$ is normalized by $CV_{tot-nr}$ , the total noise in the protein when there is no feedback. Other parameters are taken as $g_0 = 1$ , $N_x = 4$ , $V_x^2 = N_x^2 - N_x$ , $b = 0$ , $S = 1$ , $d_x = 0.04$ . The response time $T_z$ is assumed to be much larger than $T_{nr}$ . . . . .	93
4.8	A synthetic auto-regulatory gene network where the feedback strength is manipulated by adding aTc. This figure is taken from [11]. . . . .	94
4.9	Two different transcriptional responses $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$ that lead to the same intrinsic noise in the protein. $\mathbf{x}_1^*$ and $\mathbf{x}_2^*$ represents the steady-state average number of protein molecules when the transcriptional response is given by $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$ , respectively. . . . .	97
5.1	The normalized fitness $\mathbf{F}$ as a function of the feedback strength $a$ when $M = 1$ . When extrinsic noise is absent the fitness decreases with increasing $a$ , but in the presence of significant extrinsic noise, it increases. Other parameters taken as $\gamma = 1/3$ , $\kappa = 0$ , $g_0 = 1$ , $N_x = 1$ , $V_x = 0$ and $d_x = 0.01$ . . . . .	109
5.2	A plot of the evolvability of the negative feedback as a function of the Hill coefficient $M$ and normalized threshold $\gamma$ when $\kappa \ll 1$ . . . . .	111
5.3	A plot of the evolvability of the negative feedback as a function of the Hill coefficient $M$ and normalized threshold $\gamma$ when $\kappa \gg 1$ . . . . .	112

6.1	A gene activation cascade where gene $GeneX_1$ expresses protein $X_1$ . This protein then activates gene $GeneX_2$ to make $X_2$ which then goes on to activate gene $GeneX_3$ . . . . .	117
6.2	$CV_{X_i}^2$ at the $i^{th}$ stage of the gene cascade when all genes have per stage magnification of a) $G = 5$ and b) $G = 1.5$ . Other parameters taken as $r = d = \lambda = 1$ . . . . .	120
B.1	A model for gene expression with transcription and translation. . . . .	143



# Chapter 1

## Moment closure for stochastic chemical kinetics

The time evolution of a spatially homogeneous mixture of chemically reacting molecules is often modeled using a stochastic formulation, which takes into account the inherent randomness of thermal molecular motion. This formulation is superior to the traditional deterministic formulation of chemical kinetics and is motivated by complex reactions inside living cells, where small populations of key reactants can set the stage for significant stochastic effects [28, 4, 17, 58, 54].

In the stochastic formulation, the time evolution of the system is described by a single equation for a probability function, where time and species populations appear as independent variables, called the Chemical Master equation (CME) [29]. As analytical solutions to this equation are impossible (except for highly idealized cases, see [29] for examples), Monte Carlo simulation techniques [14, 15] and the finite state projection algorithm [30] are generally used to approximate the solution of the CME.

Since one is often interested in only the first and second order moments for the number of molecules of the different species involved, much effort can be saved by applying approximate methods to compute these low-order moments, without actually having to solve for the probability density function. Various such approximate methods have been developed, for example, the linear noise approximation [59, 21] and mass fluctuation kinetics, which assumes that the third order *centered* moments are equal to zero [16].

In this chapter, an alternative method for estimating low-order statistical moments is introduced for a general class of elementary chemical reactions involving  $n$  chemical species and  $K$  reactions. To derive differential equations for the time evolution of moments of the populations, the set of chemical reactions is first modeled as a Stochastic Hybrid System (SHS) with state  $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T$ , where  $\mathbf{x}_j$  denotes the population of specie  $X_j$  involved in the chemical reaction. In order to fit the framework of our problem, this SHS has trivial continuous dynamics  $\dot{\mathbf{x}} = 0$ , with reset maps and transitional intensities defined by the stoichiometry and the propensity function of the reactions, respectively. In essence, if no reaction takes place, the population of species remain constant and whenever a reaction takes place, the corresponding reset map is “activated” and the population is reset according to the stoichiometry of the reaction. Details for the stochastic modeling of chemical reactions are presented in Section 1.1. The time evolution of the moments is then obtained using results from the SHS literature [10, 18].

In Section 1.1 we construct a vector  $\mu$  containing all moments of  $\mathbf{x}$  up to order  $\mathbf{R}$ , which we refer to as an  $\mathbf{R}^{th}$  order truncation. It is shown that  $\mu$ ’s time derivative, given

by

$$\dot{\mu} = \hat{\mathbf{a}} + A\mu + B\bar{\mu} \quad (1.1)$$

for some vector  $\hat{\mathbf{a}}$  and matrices  $A$  and  $B$ , is not closed, in the sense that it depends on  $\bar{\mu}$  which is a vector containing moments of  $\mathbf{x}$  of order  $\mathbf{R} + 1$ . This assertion is based on the assumption that there is at least one bi-molecular reaction (i.e., has two reactants) in the set of chemical reactions considered. For analysis purposes, we close the above system by approximating each element of  $\bar{\mu}$  as a nonlinear function  $\varphi(\mu)$  of moments up to order  $\mathbf{R}$ . This procedure is commonly referred to as *moment closure*. We call  $\varphi(\mu)$  the *moment closure function* for the corresponding element of  $\bar{\mu}$ . We denote the state of the closed system by  $\mathbf{v}$ , which should be viewed as an approximation for the vector  $\mu$ . The dynamics for  $\mathbf{v}$  are given by

$$\dot{\mathbf{v}} = \hat{\mathbf{a}} + A\mathbf{v} + B\bar{\varphi}(\mathbf{v}), \quad (1.2)$$

where  $\bar{\varphi}$  is a column vector of moment closure functions  $\varphi$ . The above closed dynamics is referred to as the *truncated moment dynamics*.

A general procedure to construct these moment closure functions is outlined. This procedure is based on first assuming a certain *separable form* for each element  $\varphi$  of  $\bar{\varphi}$  and then matching time derivatives of  $\mu$  and  $\mathbf{v}$  in (1.1) and (1.2) respectively, at some initial time  $t_0$ , for a deterministic initial condition  $\mathbf{x}(t_0) = \mathbf{x}_0$  with probability one. This procedure uniquely determines the moment closure functions  $\varphi$  and explicit formulas to construct them are provided in Section 1.2. We refer to this moment closure procedure as the *Separable Derivative-Matching (SDM) moment closure*.

One of the key results of this section is that for SDM moment closure we have

$$\left. \frac{d^i \mu(t)}{dt^i} \right|_{t=t_0} = \left. \frac{d^i \mathbf{v}(t)}{dt^i} \right|_{t=t_0} + \varepsilon_i(\mathbf{x}_0), \quad \forall i \geq 1 \quad (1.3)$$

for deterministic initial conditions, i.e.,  $\mathbf{x}(t_0) = \mathbf{x}_0$  with probability one. In (1.3), each element of vector  $\frac{d^i \mu(t)}{dt^i} \big|_{t=t_0}$  is a polynomial in  $\mathbf{x}_0$ , whose degree is larger than the degree of the corresponding polynomial element of the error vector  $\varepsilon_i(\mathbf{x}_0)$  by  $\mathbf{R}$  whenever that element of the error vector is non-zero. Thus, with increasing  $\mathbf{R}$ , the truncated moment dynamics  $\mathbf{v}(t)$  should provide more accurate approximations to the lower order moments  $\mu(t)$ .

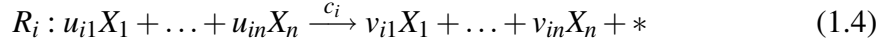
A striking feature of the SDM moment closure functions is that they are independent of the reaction parameters (reaction rates and stoichiometry) and moreover the dependence of higher-order moment on lower order ones is consistent with  $\mathbf{x}$  being jointly *lognormally distributed* in spite of the fact that the procedure used to construct  $\varphi$  does not make any assumption on the distribution of the population.

To illustrate the applicability of our results, we construct truncated moment dynamics for a set of chemical reactions motivated by a gene cascade network. It is shown that truncated moment dynamics based on a second order truncation provides good estimates of the first and second order moments as compared with the actual moments obtained using Monte Carlo simulations (Gillespie's SSA [14]). As claimed, performance improves with a third order truncation, which provides more accurate estimates of the actual moment dynamics.

The moment closure procedure proposed in this chapter is also compared to an alternative procedure where moment closure is performed by setting the third order *centered* moments equal to zero [16]. This later procedure is referred to as the *normal moment closure*. It is shown that for large populations, the normal moment closure provides better estimates of the statistical moments than the SDM moment closure. However, when the population size is small the SDM moment closure outperforms the normal moment closure.

## 1.1 Moment dynamics of chemically reacting systems

Consider a system of  $n$  species  $X_j$ ,  $j \in \{1, \dots, n\}$  inside a fixed volume  $V$  involved in  $K$  reactions of the form



for all  $i \in \{1, \dots, K\}$ , where  $u_{ij} \in \mathbb{N}_{\geq 0}$  is the stoichiometry associated with the  $j^{\text{th}}$  reactant in the  $i^{\text{th}}$  reaction and  $v_{ij} \in \mathbb{N}_{\geq 0}$  is the stoichiometry associated with the  $j^{\text{th}}$  product in the  $i^{\text{th}}$  reaction, and  $*$  represents products other than the species  $X_j$ . As all chemical reactions occur in a series of elementary reactions, which are generally uni- or bi-molecular, we assume

$$u_{i1} + \dots + u_{in} \leq 2, \quad \forall i \in \{1, \dots, K\}, \quad (1.5)$$

[65], and hence, we only allow reactions which have the form given in the first column of Table 1.1. The parameter  $c_i$  characterizes the  $i^{\text{th}}$  reaction  $R_i$  and defines the probability that this reaction will take place in the next “infinitesimal” time interval  $(t, t + dt]$ . This probability is given by  $c_i h_i dt$  where  $h_i$  is the number of distinct molecular reactant combinations for the reaction  $R_i$  present in the volume  $V$  at time  $t$  and  $c_i dt$  is the probability that a particular reactant combination of  $R_i$  will actually react on  $(t, t + dt]$ . The product  $c_i h_i$  is called the *propensity function* of reaction  $R_i$ . The number  $h_i$  depends both on the reactants stoichiometry  $u_{ij}$  in  $R_i$  and on the number of reactant molecules in  $V$ . Table 1.1 shows the form of  $h_i$  for standard elementary reactions [14]. In this table and in the sequel, we denote by  $\mathbf{x}_j$ , the number of molecules of the species  $X_j$  in the volume  $V$  and  $\mathbf{x} := [\mathbf{x}_1, \dots, \mathbf{x}_n]^T \in \mathbb{R}^n$ . The reaction parameter  $c_i$  is related to the reaction rate  $k_i$  in the deterministic formulation of chemical kinetics by the formulas shown in the right-most column of Table 1.1.

Table 1.1.  $h_i(\mathbf{x})$  and  $c_i$  for different elementary reactions.

Reaction $R_i$	$h_i(\mathbf{x})$	$c_i$
$X_j \longrightarrow *$	$\mathbf{x}_j$	$k_i$
$X_j + X_\ell \longrightarrow *$ ( $\ell \neq j$ )	$\mathbf{x}_j \mathbf{x}_\ell$	$\frac{k_i}{V}$
$2X_j \longrightarrow *$	$\frac{1}{2} \mathbf{x}_j (\mathbf{x}_j - 1)$	$\frac{2k_i}{V}$

To model the time evolution of the number of molecules  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ , a special class of Stochastic Hybrid Systems (SHS) was introduced in [19]. More specifically, to fit the framework of our problem, these systems are characterized by trivial dynamics

$$\dot{\mathbf{x}} = 0, \quad \mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T, \quad (1.6)$$

a family of  $K$  reset maps

$$\mathbf{x} = \phi_i(\mathbf{x}^-), \quad \phi_i : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad (1.7)$$

and a corresponding family of  $K$  transition intensities

$$\lambda_i(\mathbf{x}), \quad \lambda_i : \mathbb{R}^n \rightarrow [0, \infty) \quad (1.8)$$

for all  $i \in \{1, \dots, K\}$ . Each of the reset maps  $\phi_i(\mathbf{x})$ , and corresponding transition intensities  $\lambda_i(\mathbf{x})$  are uniquely defined by the  $i^{th}$  reaction and given by

$$\mathbf{x} \mapsto \phi_i(\mathbf{x}) = \begin{bmatrix} \mathbf{x}_1 - u_{i1} + v_{i1} \\ \mathbf{x}_2 - u_{i2} + v_{i2} \\ \vdots \\ \mathbf{x}_n - u_{in} + v_{in} \end{bmatrix}, \quad \lambda_i(\mathbf{x}) = c_i h_i(\mathbf{x}) \quad (1.9)$$

for all  $i \in \{1, \dots, K\}$ . In essence, if no reaction takes place, the state remains constant and whenever the  $i^{th}$  reaction takes place, the reset map  $\phi_i(\mathbf{x})$  is “activated” and the

state  $\mathbf{x}$  is reset according to (1.9), furthermore, the probability of the activation taking place in an “infinitesimal” time interval  $(t, t + dt]$  is  $\lambda_i(\mathbf{x})dt$ .

Given a vector  $\mathbf{m} = (m_1, m_2, \dots, m_n) \in \mathbb{N}_{\geq 0}^n$  of  $n$  non-negative integers, we define the (*uncentered*) moment of  $\mathbf{x}$  associated with  $\mathbf{m}$  to be

$$\mu_{(\mathbf{m})}(t) := \mathbf{E}[\mathbf{x}^{(\mathbf{m})}(t)], \quad \forall t \geq 0 \quad (1.10)$$

where  $\mathbf{E}$  stands for the expected value and

$$\mathbf{x}^{(\mathbf{m})} := \mathbf{x}_1^{m_1} \mathbf{x}_2^{m_2} \dots \mathbf{x}_n^{m_n}. \quad (1.11)$$

The sum  $\sum_{j=1}^n m_j$  is called the *order of the moment*. For example, consider a system of reactions with two species ( $n = 2$ ) and  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2]^T$ . Then, the first order moments are given by

$$\mu_{(1,0)} = \mathbf{E}[\mathbf{x}_1], \quad \mu_{(0,1)} = \mathbf{E}[\mathbf{x}_2], \quad (1.12)$$

the second order moments are given by

$$\mu_{(2,0)} = \mathbf{E}[\mathbf{x}_1^2], \quad \mu_{(0,2)} = \mathbf{E}[\mathbf{x}_2^2], \quad \mu_{(1,1)} = \mathbf{E}[\mathbf{x}_1 \mathbf{x}_2], \quad (1.13)$$

and so on. In the sequel, when we simply say a “moment of  $\mathbf{x}$ ,” we refer to an *uncentered* moment of  $\mathbf{x}$ . The time evolution of the moments is given by the following Theorem (see Appendix A for details).

**Theorem 1.1** *Let the vector*

$$\boldsymbol{\mu} = [\mu_{(\mathbf{m}_1)}, \dots, \mu_{(\mathbf{m}_k)}]^T \in \mathbb{R}^k, \quad \mathbf{m}_p \in \mathbb{N}_{\geq 0}^n, \quad \forall p \in \{1, \dots, k\} \quad (1.14)$$

*contain all the moments of  $\mathbf{x}$  of order up to  $\mathbf{R}$ . Let there exists at least one reaction in (1.4) which has two reactants, then the time derivative of  $\boldsymbol{\mu}$  is given by*

$$\dot{\boldsymbol{\mu}} = \hat{\mathbf{a}} + A\boldsymbol{\mu} + B\bar{\boldsymbol{\mu}}, \quad (1.15)$$

for an appropriate vector  $\hat{\mathbf{a}}$ , matrices  $A$  and  $B$ , and  $\bar{\mu} \in \mathbb{R}^r$  is a vector of moments of order  $\mathbf{R} + 1$ .

In the sequel, we refer to  $\mathbf{R}$  as the order truncation. One can see from the above Theorem that the dynamics of vector  $\mu$  is not closed, in the sense that it depends on higher order moments in  $\bar{\mu}$ . We approximate (1.15) by a nonlinear system of the form

$$\dot{\mathbf{v}} = \hat{\mathbf{a}} + A\mathbf{v} + B\bar{\varphi}(\mathbf{v}), \quad \mathbf{v} = [v_{(\mathbf{m}_1)}, \dots, v_{(\mathbf{m}_k)}]^T \quad (1.16)$$

where the map  $\bar{\varphi} : \mathbb{R}^k \rightarrow \mathbb{R}^r$  should be chosen so as to keep  $\mathbf{v}(t)$  close to  $\mu(t)$ . This procedure is commonly referred to as *moment closure*. We call (1.16) the *truncated moment dynamics* and each element  $\varphi_{(\bar{\mathbf{m}})}(\mu)$  of  $\bar{\varphi}(\mu)$  the *moment closure function* for the corresponding element  $\mu_{(\bar{\mathbf{m}})}$  in  $\bar{\mu}$ .

When a sufficiently large number of derivatives of  $\mu(t)$  and  $\mathbf{v}(t)$  match point-wise, then, the difference between solutions to (1.15) and (1.16) remains close on a given compact time interval, this follows from a Taylor series approximation argument. To be more precise, for each  $\delta > 0$  and  $T \in \mathbb{R}$ , there exists an integer  $N$ , sufficiently large, for which the following result holds: Assuming that for some initial time  $t_0$ ,

$$\mu(t_0) = \mathbf{v}(t_0) \Rightarrow \frac{d^i \mu(t)}{dt^i} \Big|_{t=t_0} = \frac{d^i \mathbf{v}(t)}{dt^i} \Big|_{t=t_0}, \quad \forall i \in \{1, 2, \dots, N\} \quad (1.17)$$

where  $\frac{d^i \mu(t)}{dt^i} \Big|_{t=t_0}$  and  $\frac{d^i \mathbf{v}(t)}{dt^i} \Big|_{t=t_0}$  represent the  $i^{th}$  time derivative of  $\mu(t)$  and  $\mathbf{v}(t)$  along the trajectories of system (1.15) and (1.16), respectively at an initial time  $t_0$ . Then,

$$\|\mu(t) - \mathbf{v}(t)\| \leq \delta, \quad \forall t \in [t_0, T], \quad (1.18)$$

along solutions of (1.15) and (1.16). Note that for the above argument to hold there is an implicit assumption that both  $\mu(t)$  and  $\mathbf{v}(t)$  are analytical functions, i.e., can be written as a convergent power series about  $t = t_0$ .



## 1.2 SDM moment closure procedure

In this section we construct truncated moment dynamics so that the condition (1.17) holds approximately. After replacing (1.15) and (1.16) in (1.17) we obtain a PDE on  $\bar{\varphi}$ . We seek solutions  $\bar{\varphi}$  to this PDE for which each entry  $\varphi_{(\mathbf{m})}$  has the following separable form

$$\varphi_{(\mathbf{m})}(\mathbf{v}) = \prod_{p=1}^k \left( v_{(\mathbf{m}_p)} \right)^{\gamma_p}, \quad (1.19)$$

where  $\gamma_p$  are appropriately chosen constants.

### 1.2.1 Derivative matching for a generator

In general, it will not be possible to find constants  $\gamma_p$  such that (1.17) holds for any arbitrary initial conditions. To be more precise, denoting by  $\mu_{\infty}(t)$  an infinite vector containing all moments of  $\mathbf{x}$  and by  $\Omega_{\infty}$  the (convex) set of all possible values that  $\mu_{\infty}$  can take, it will not be possible to make (1.17) hold  $\forall \mu_{\infty}(t_0) \in \Omega_{\infty}$ . Instead, we will simply require (1.17) to hold for vectors  $\mu_{\infty}(t_0)$  belonging to a set of vectors that generate  $\Omega_{\infty}$  by convex combinations. It is not hard to see that the set of deterministic distributions forms a natural generator for  $\Omega_{\infty}$ . For example, with two species ( $n = 2$ ) with populations  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , denoting by  $P_{x_1 x_2}(t)$  the probability of having  $\mathbf{x}_1(t) = x_1$ ,

$\mathbf{x}_2(t) = x_2$ , the infinite vector  $\mu_\infty$  can be expressed as

$$\mu_\infty(t) = \begin{bmatrix} \mathbf{E}[\mathbf{x}_1(t)] \\ \mathbf{E}[\mathbf{x}_2(t)] \\ \mathbf{E}[\mathbf{x}_1^2(t)] \\ \mathbf{E}[\mathbf{x}_2^2(t)] \\ \mathbf{E}[\mathbf{x}_1(t)\mathbf{x}_2(t)] \\ \vdots \end{bmatrix} = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \\ x_2^2 \\ x_1 x_2 \\ \vdots \end{bmatrix} P_{x_1 x_2}(t). \quad (1.20)$$

Hence, the set of infinite vectors

$$\mathbf{D} = \{[x_1, x_2, x_1^2, x_2^2, x_1 x_2, \dots]^T : x_1, x_2 \geq 0\} \quad (1.21)$$

which corresponds to deterministic distributions, i.e.,  $\mathbf{x}_1(t) = x_1$ ,  $\mathbf{x}_2(t) = x_2$  with probability one, generated  $\Omega_\infty$  by convex combinations. In the sequel, we seek for constants  $\gamma_p$  for which (1.17) holds for every vector  $\mu_\infty(t_0)$  belonging to  $\mathbf{D}$ , i.e., for the class of deterministic initial conditions.

However, often it is still not possible to find  $\gamma_p$  for which (1.17) holds exactly in  $\mathbf{D}$ . We will therefore relax this condition and simply demand the following

$$\mu(t_0) = \mathbf{v}(t_0) \Rightarrow \frac{d^i \mu(t)}{dt^i} \Big|_{t=t_0} = \frac{d^i \mathbf{v}(t)}{dt^i} \Big|_{t=t_0} + \mathbf{E}[\varepsilon_i(\mathbf{x}(t_0))], \quad (1.22)$$

$\forall i \in \{1, 2, \dots, N\}$ , where each element of  $\varepsilon_i(\mathbf{x}(t_0))$  is a polynomial in  $\mathbf{x}(t_0)$ . One can think of (1.22) as an approximation to (1.17) which will be valid as long as the moments in  $\frac{d^i \mu(t)}{dt^i} \Big|_{t=t_0}$  dominate over  $\mathbf{E}[\varepsilon_i(\mathbf{x}(t_0))]$ .

Before stating our main result, for given vectors  $\hat{\mathbf{m}} = (\hat{m}_1, \dots, \hat{m}_n) \in \mathbb{N}_{\geq 0}^n$  and  $\check{\mathbf{m}} =$

$(\check{m}_1, \dots, \check{m}_n) \in \mathbb{N}_{\geq 0}^n$ , we define the scalar<sup>1</sup>

$$\mathbf{C}_{(\check{\mathbf{m}})}^{(\hat{\mathbf{m}})} := \mathbf{C}_{\check{m}_1}^{\hat{m}_1} \mathbf{C}_{\check{m}_2}^{\hat{m}_2} \dots \mathbf{C}_{\check{m}_n}^{\hat{m}_n}. \quad (1.23)$$

**Theorem 1.2** Assume that for each element  $\mu_{(\bar{\mathbf{m}})}$  of  $\bar{\mu}$ , the corresponding moment closure function  $\varphi_{(\bar{\mathbf{m}})}$  in  $\bar{\varphi}$  is chosen according to (1.19) with  $\gamma = (\gamma_1, \dots, \gamma_k)$  taken as the unique solution of the following system of linear equations

$$\mathbf{C}_{(\mathbf{m}_s)}^{(\bar{\mathbf{m}})} = \sum_{p=1}^k \gamma_p \mathbf{C}_{(\mathbf{m}_s)}^{(\mathbf{m}_p)}, \quad \forall s = \{1, \dots, k\}. \quad (1.24)$$

Then, with initial conditions  $\mathbf{x}(t_0) = \mathbf{x}_0 \in \mathbb{R}^n$  with probability one, we have that

$$\mu(t_0) = \mathbf{v}(t_0) \Rightarrow \frac{d\mu(t)}{dt} \Big|_{t=t_0} = \frac{d\mathbf{v}(t)}{dt} \Big|_{t=t_0} \quad (1.25a)$$

$$\Rightarrow \frac{d^2\mu(t)}{dt^2} \Big|_{t=t_0} = \frac{d^2\mathbf{v}(t)}{dt^2} \Big|_{t=t_0} + \varepsilon_2(\mathbf{x}_0) \quad (1.25b)$$

where the last  $\bar{n}$  elements of  $\varepsilon_2(\mathbf{x}_0)$  are polynomials in  $\mathbf{x}_0$  of degree 2 and all other elements are zero. Here,  $\bar{n} = k - n^* + 1$ , where  $n^*$  denotes the row in the vector  $\mu$  from where the  $\mathbf{R}^{th}$  order moments of  $\mathbf{x}$  start appearing.

As the last  $\bar{n}$  elements of  $\mu(t)$  contain moments of order  $\mathbf{R}$ , we have that the last  $\bar{n}$  elements of the vector  $\frac{d^2\mu(t)}{dt^2} \Big|_{t=t_0}$  are polynomials in  $\mathbf{x}_0$  of degree  $\mathbf{R} + 2$ . Thus, we have for sufficiently large  $\|\mathbf{x}_0\|$  that

$$\frac{\varepsilon_2^p(\mathbf{x}_0)}{\frac{d^2\mu(\mathbf{m}_p)(t)}{dt^2} \Big|_{t=t_0}} = \begin{cases} 0 & \text{if } p < n^* \\ O(\|\mathbf{x}_0\|^{-\mathbf{R}}) & \text{if } p \geq n^* \end{cases} \quad (1.26)$$

---

<sup>1</sup> $\mathbf{C}_h^\ell$  is defined as follows:  $\forall \ell, h \in \mathbb{N}_{\geq 0}$

$$\mathbf{C}_h^\ell = \begin{cases} \frac{\ell!}{(\ell-h)!h!}, & \ell \geq h \\ 0, & \ell < h \end{cases}$$

where  $\ell!$  denotes the factorial of  $\ell$ . The factorial of zero is defined as one. Hence by this definition  $\mathbf{C}_0^0 = 1$ .

where  $\varepsilon_2^p(\mathbf{x}_0)$  and  $\mu_{(\mathbf{m}_p)}(t)$  denote the  $p^{th}$  element of the vector  $\varepsilon_2(\mathbf{x}_0)$  and  $\mu(t)$ , respectively, and  $O$  denotes the order of magnitude. This result shows that the error  $\varepsilon_2$  in the derivative-matching can be reduced by increasing the order of truncation  $\mathbf{R}$ .

From the proof of Theorem 1.2 one can see that the functions  $h_i(\mathbf{x})$  only determine the error  $\varepsilon_2(\mathbf{x}_0)$  and not the constants  $\gamma_p$ . Thus, Theorem 1.2 can also be used to compute moment closure functions when the reactions are non-elementary and  $h_i(\mathbf{x})$  is a polynomial of arbitrary degree. The only difference would be that the higher order moments to be approximated can be of order  $\mathbf{R} + 1$  or higher.

It can be verified that with  $\gamma_p$  chosen as in Theorem 1.2, moment closure functions so obtained also match derivatives of order higher than 2 in (1.25b) with small errors. Using symbolic manipulation in *Mathematica*, for  $n \in \{1, 2, 3\}$  and  $i \in \{3, 4, 5\}$  one can verify that the degree of the polynomial elements of  $\frac{d^i \mu(t)}{dt^i} \big|_{t=t_0}$  are larger than the degree of the corresponding polynomial elements in the error vector  $\varepsilon_i(\mathbf{x}_0) = \frac{d^i \mu(t)}{dt^i} \big|_{t=t_0} - \frac{d^i v(t)}{dt^i} \big|_{t=t_0}$  by  $\mathbf{R}$  whenever that element in the error vector is non-zero. We conjecture that this is true  $\forall n \in \mathbb{N}$  and  $\forall i \in \mathbb{N}$  but we only verified it for  $n$  up to 3 and  $i$  up to 5. Hence, with increasing  $\mathbf{R}$ , the truncated moment dynamics  $v(t)$  should provide a more accurate approximations to the lower order moments  $\mu(t)$ . In the sequel we refer to moment closure functions obtained from (1.19) and (1.24) as *Separable Derivative-Matching (SDM) moment closure functions*.

### 1.2.2 Proof of Theorem 1.2

Let the vector  $\mathbf{m}_p \in \mathbf{O}_m$  denote that  $\mu_{(\mathbf{m}_p)}$  is a moment of order  $m$ . We assume that the moments in vector  $\mu$  are arranged in increasing order, i.e.,

$$\mathbf{m}_p \in \mathbf{O}_1, \quad \forall p \in \{1, \dots, n\} \quad (1.27a)$$

$$\mathbf{m}_p \in \mathbf{O}_2, \quad \forall p \in \{n+1, \dots, n+n(n+1)/2\} \quad (1.27b)$$

and so on. Moreover, the vector  $\mathbf{m}_p \in \mathbf{O}_1$  has 1 in the  $p^{th}$  position and all other entries are zero. From (1.15) and (1.16) we have that

$$\left. \frac{d\mu(t)}{dt} \right|_{t=t_0} - \left. \frac{dv(t)}{dt} \right|_{t=t_0} = A(\mu(t_0) - v(t_0)) + B(\bar{\mu}(t_0) - \bar{\varphi}(v(t_0))) \quad (1.28a)$$

$$\begin{aligned} \left. \frac{d^2\mu(t)}{dt^2} \right|_{t=t_0} - \left. \frac{d^2v(t)}{dt^2} \right|_{t=t_0} = & A \left( \left. \frac{d\mu(t)}{dt} \right|_{t=t_0} - \left. \frac{dv(t)}{dt} \right|_{t=t_0} \right) + \\ & B \left( \left. \frac{d\bar{\mu}(t)}{dt} \right|_{t=t_0} - \left. \frac{d\bar{\varphi}(v(t))}{dt} \right|_{t=t_0} \right). \end{aligned} \quad (1.28b)$$

Using  $\mu(t_0) = v(t_0)$  and the fact that all elements in the first  $n^* - 1$  rows of matrix  $B$  are zero (see Appendix A), it is sufficient to prove the following : For each element  $\mu_{(\bar{\mathbf{m}})}$  of  $\bar{\mu}$  and its corresponding moment closure function  $\varphi_{(\bar{\mathbf{m}})}(\mu)$  we have

$$\mu_{(\bar{\mathbf{m}})}(t_0) = \varphi_{(\bar{\mathbf{m}})}(\mu(t_0)) \quad (1.29a)$$

$$\left. \frac{d\mu_{(\bar{\mathbf{m}})}(t)}{dt} \right|_{t=t_0} = \left. \frac{d\varphi_{(\bar{\mathbf{m}})}(\mu(t))}{dt} \right|_{t=t_0} + \bar{\varepsilon}(\mathbf{x}_0) \quad (1.29b)$$

where the scalar  $\bar{\varepsilon}(\mathbf{x}_0)$  is polynomial in  $\mathbf{x}_0$  of degree 2.

We first prove (1.29a). For the initial conditions  $\mathbf{x}(t_0) = \mathbf{x}_0$  with probability one and using (1.19) we have

$$\mu_{(\bar{\mathbf{m}})}(t_0) = \mathbf{x}_0^{(\bar{\mathbf{m}})}, \quad (1.30a)$$

$$\varphi_{(\bar{\mathbf{m}})}(\mu(t_0)) = \prod_{p=1}^k \left( \mathbf{x}_0^{(\mathbf{m}_p)} \right)^{\gamma_p} = \mathbf{x}_0^{(\sum_{p=1}^k \gamma_p \mathbf{m}_p)}. \quad (1.30b)$$

Using (1.23) and the fact that the vectors  $\mathbf{m}_p \in \mathbf{O}_1$ ,  $p \in \{1, \dots, n\}$  have 1 in the  $p^{th}$  position and all other entries zero, we have that

$$\bar{\mathbf{m}} = \left( \mathbf{C}_{(\mathbf{m}_1)}^{(\bar{\mathbf{m}})}, \dots, \mathbf{C}_{(\mathbf{m}_n)}^{(\bar{\mathbf{m}})} \right), \quad \mathbf{m}_p = \left( \mathbf{C}_{(\mathbf{m}_1)}^{(\mathbf{m}_p)}, \dots, \mathbf{C}_{(\mathbf{m}_n)}^{(\mathbf{m}_p)} \right), \quad \forall p \in \{1, \dots, k\}. \quad (1.31)$$

From the above equalities and using (1.24) for  $s \in \{1, \dots, n\}$  we have that  $\bar{\mathbf{m}} = \sum_{p=1}^k \gamma_p \mathbf{m}_p$ .

Hence from (1.30a)-(1.30b), equality (1.29a) holds.

Our next goal is to prove (1.29b). We have from Appendix A that the time derivative of a moment  $\mu_{(\mathbf{m})}$  of order  $m$  is given by (A.3) where from (A.5a)

$$(\mathbf{L}\psi)(\mathbf{x}) = \sum_{i=1}^K c_i h_i(\mathbf{x}) \left\{ \left[ \prod_{j=1}^n \sum_{q=0}^{m_j} \mathbf{C}_q^{m_j} \mathbf{x}_j^{m_j-q} a_{ij}^q \right] - \mathbf{x}^{(\mathbf{m})} \right\} := \sum_{i=1}^K c_i h_i(\mathbf{x}) g_i(\mathbf{x}). \quad (1.32)$$

Note that the polynomial  $g_i(\mathbf{x})$  is a polynomial of order  $m-1$ . The monomials that contribute to terms of degree  $m-1$  in  $g_i(\mathbf{x})$  are

$$\sum_{j=1}^n \mathbf{C}_1^{m_j} \frac{\mathbf{x}^{(\mathbf{m})}}{\mathbf{x}_j} a_{ij} = \sum_{p=1}^n \mathbf{C}_{(\mathbf{m}_p)}^{(\mathbf{m})} \mathbf{x}^{(\mathbf{m}-\mathbf{m}_p)} \mathbf{a}_i^{(\mathbf{m}_p)} = \sum_{\forall \mathbf{m}_p \in \mathbf{O}_1} \mathbf{C}_{(\mathbf{m}_p)}^{(\mathbf{m})} \mathbf{x}^{(\mathbf{m}-\mathbf{m}_p)} \mathbf{a}_i^{(\mathbf{m}_p)} \quad (1.33)$$

where  $\mathbf{a}_i = [a_{i1}, \dots, a_{in}]^T$ . Similarly the monomials that contributes to terms of degree  $m-2$  are

$$\sum_{j=1}^n \mathbf{C}_2^{m_j} \frac{\mathbf{x}^{(\mathbf{m})}}{\mathbf{x}_j^2} a_{ij}^2 + \sum_{j=1}^n \sum_{u=1 (u \neq j)}^n \mathbf{C}_1^{m_j} \mathbf{C}_1^{m_u} \frac{\mathbf{x}^{(\mathbf{m})}}{\mathbf{x}_j \mathbf{x}_u} a_{ij} a_{iu} = \sum_{\forall \mathbf{m}_p \in \mathbf{O}_2} \mathbf{C}_{(\mathbf{m}_p)}^{(\mathbf{m})} \mathbf{x}^{(\mathbf{m}-\mathbf{m}_p)} \mathbf{a}_i^{(\mathbf{m}_p)}. \quad (1.34)$$

Thus the term

$$\sum_{l=1}^{\mathbf{R}} \sum_{\forall \mathbf{m}_p \in \mathbf{O}_l} \mathbf{C}_{(\mathbf{m}_p)}^{(\mathbf{m})} \mathbf{x}^{(\mathbf{m}-\mathbf{m}_p)} \mathbf{a}_i^{(\mathbf{m}_p)} = \sum_{s=1}^k \mathbf{C}_{(\mathbf{m}_s)}^{(\mathbf{m})} \mathbf{x}^{(\mathbf{m}-\mathbf{m}_s)} \mathbf{a}_i^{(\mathbf{m}_s)} \quad (1.35)$$

represents all terms that contribute to monomials of degree  $m-1$  to  $m-\mathbf{R}$  in  $g_i(\mathbf{x})$ .

Therefore we can write

$$\frac{d\mu_{(\mathbf{m})}(t)}{dt} = \mathbf{E}[(\mathbf{L}\psi)(\mathbf{x})], \quad (\mathbf{L}\psi)(\mathbf{x}) = \sum_{i=1}^K c_i h_i(\mathbf{x}) g_i(\mathbf{x}) \quad (1.36a)$$

$$g_i(\mathbf{x}) = \sum_{s=1}^k \mathbf{C}_{(\mathbf{m}_s)}^{(\mathbf{m})} \mathbf{x}^{(\mathbf{m}-\mathbf{m}_s)} \mathbf{a}_i^{(\mathbf{m}_s)} + P_{m-1-\mathbf{R}}(\mathbf{x}) \quad (1.36b)$$

where  $P_{m-1-\mathbf{R}}(\mathbf{x})$  is zero if  $m-1-\mathbf{R} < 0$ , constant if  $m-1-\mathbf{R} = 0$  and a polynomial in  $\mathbf{x}$  of degree  $m-1-\mathbf{R}$  otherwise. As  $\mu_{(\tilde{\mathbf{m}})}$  is a moment of order  $\mathbf{R}+1$  (hence  $m = \mathbf{R}+1$ ) we have from (1.36)

$$\frac{d\mu_{(\tilde{\mathbf{m}})}(t)}{dt}\Big|_{t=t_0} = \sum_{i=1}^K c_i h_i(\mathbf{x}_0) \left\{ \sum_{s=1}^k \mathbf{C}_{(\mathbf{m}_s)}^{(\tilde{\mathbf{m}})} \mathbf{x}_0^{(\tilde{\mathbf{m}}-\mathbf{m}_s)} \mathbf{a}_i^{(\mathbf{m}_s)} + G_i \right\} \quad (1.37)$$

where  $G_i$  is a constant. Also from (1.19) and using (1.29a), (1.36)

$$\frac{d\varphi_{(\tilde{\mathbf{m}})}(\mu(t))}{dt}\Big|_{t=t_0} = \varphi_{(\tilde{\mathbf{m}})}(\mu(t_0)) \sum_{p=1}^k \gamma_p \frac{d\mu_{(\mathbf{m}_p)}(t_0)}{dt}\Big|_{t=t_0} / \mu_{(\mathbf{m}_p)}(t_0) \quad (1.38a)$$

$$= \sum_{p=1}^k \gamma_p \mathbf{x}_0^{(\tilde{\mathbf{m}}-\mathbf{m}_p)} \frac{d\mu_{(\mathbf{m}_p)}(t)}{dt}\Big|_{t=t_0} \quad (1.38b)$$

$$= \sum_{i=1}^K c_i h_i(\mathbf{x}_0) \left\{ \sum_{s=1}^k \sum_{p=1}^k \gamma_p \mathbf{C}_{(\mathbf{m}_s)}^{(\mathbf{m}_p)} \mathbf{x}_0^{(\tilde{\mathbf{m}}-\mathbf{m}_s)} \mathbf{a}_i^{(\mathbf{m}_s)} \right\}. \quad (1.38c)$$

Using (1.24), (1.37) and (1.38c) one can see that

$$\bar{\varepsilon}(\mathbf{x}_0) = \sum_{i=1}^K c_i h_i(\mathbf{x}_0) G_i, \quad (1.39)$$

and is a polynomial of degree 2.

### 1.2.3 SDM moment closure functions

In this section, we use (1.24) to construct moment closure functions for different higher order moments  $\mu_{(\tilde{\mathbf{m}})}$  corresponding to different number of species  $n$  and various orders of truncation  $\mathbf{R}$ .

#### Single-species reactions ( $n = 1$ )

For a  $\mathbf{R}^{th}$  order truncation we have

$$\mu = [\mu_1, \dots, \mu_k]^T := \left[ \mathbf{E}[\mathbf{x}_1], \dots, \mathbf{E}[\mathbf{x}_1^k] \right]^T, \quad \bar{\mu} = \mu_{k+1} := \mathbf{E}[\mathbf{x}_1^{k+1}] \quad (1.40)$$

in (1.15). Note that for single-species reactions the size of the vector  $\mu$  is equal to the order of truncation, i.e.,  $k = \mathbf{R}$ . Using (1.23), (1.24) reduces to

$$\mathbf{C}_s^{k+1} = \sum_{p=1}^k \gamma_p \mathbf{C}_s^p, \quad \forall s = \{1, \dots, k\}. \quad (1.41)$$

It is shown in Chapter 3, Theorem 3.1 that the unique solution to the above system of linear equations is

$$\gamma_p = (-1)^{k-p} \mathbf{C}_p^{k+1}, \quad \forall p = \{1, \dots, k\}. \quad (1.42)$$

Table 1.2 shows the corresponding moment closure functions  $\varphi_{k+1}(\mu)$  for  $\bar{\mu} = \mu_{k+1} = \mathbf{E}[\mathbf{x}_1^{k+1}]$  obtained for truncations of order  $k$  equal to 2, 3 and 4.

Table 1.2. Moment closure function  $\varphi_{k+1}(\mu)$  for  $\bar{\mu} = \mu_{k+1} = \mathbf{E}[\mathbf{x}_1^{k+1}]$  with  $k \in \{2, 3, 4\}$ .

	$k = 2$	$k = 3$	$k = 4$
$\varphi_{k+1}(\mu)$	$\left( \frac{\mathbf{E}[\mathbf{x}_1^2]}{\mathbf{E}[\mathbf{x}_1]} \right)^3$	$\frac{\mathbf{E}[\mathbf{x}_1^4]}{\mathbf{E}[\mathbf{x}_1^2]^2} \left( \frac{\mathbf{E}[\mathbf{x}_1^3]}{\mathbf{E}[\mathbf{x}_1]} \right)^4$	$\left( \frac{\mathbf{E}[\mathbf{x}_1^2]}{\mathbf{E}[\mathbf{x}_1]} \right)^{10} \left( \frac{\mathbf{E}[\mathbf{x}_1^4]}{\mathbf{E}[\mathbf{x}_1]} \right)^5$

### Two-species reactions ( $n = 2$ )

We first consider a second order truncation ( $\mathbf{R} = 2$ ), for which we have

$$\mu = [\mu_{(1,0)}, \mu_{(0,1)}, \mu_{(2,0)}, \mu_{(0,2)}, \mu_{(1,1)}]^T := [\mathbf{E}[\mathbf{x}_1], \mathbf{E}[\mathbf{x}_2], \mathbf{E}[\mathbf{x}_1^2], \mathbf{E}[\mathbf{x}_2^2], \mathbf{E}[\mathbf{x}_1 \mathbf{x}_2]]^T. \quad (1.43)$$

Recall that for  $\mathbf{R} = 2$ , the vector  $\bar{\mu}$  in (1.15) contains third order moments of  $\mathbf{x}$ . Consider an element  $\mu_{(2,1)} := \mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2]$  of  $\bar{\mu}$  and let the corresponding moment closure



function be given by

$$\varphi_{(2,1)}(\mu) = (\mu_{(1,0)})^{\gamma_1} (\mu_{(0,1)})^{\gamma_2} (\mu_{(2,0)})^{\gamma_3} (\mu_{(0,2)})^{\gamma_4} (\mu_{(1,1)})^{\gamma_5}. \quad (1.44)$$

Its not hard to verify that the equations in (1.24) reduce to

$$2 = \gamma_1 + 2\gamma_3 + \gamma_5, \quad 1 = \gamma_2 + 2\gamma_4 + \gamma_5, \quad 1 = \gamma_3, \quad 0 = \gamma_4, \quad 2 = \gamma_5 \quad (1.45)$$

leading to the following moment closure function for  $\mu_{(2,1)} = \mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2]$ :

$$\varphi_{(2,1)}(\mu) = \left( \frac{\mu_{(2,0)}}{\mu_{(0,1)}} \right) \left( \frac{\mu_{(1,1)}}{\mu_{(1,0)}} \right)^2 = \left( \frac{\mathbf{E}[\mathbf{x}_1^2]}{\mathbf{E}[\mathbf{x}_2]} \right) \left( \frac{\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2]}{\mathbf{E}[\mathbf{x}_1]} \right)^2. \quad (1.46)$$

Repeating the above analysis for different third-order moments, we obtain the moment closure functions in Table 1.3. Note that the moment closure function for  $\mu_{(1,2)} =$

Table 1.3. Moment closure function  $\varphi_{(\bar{\mathbf{m}})}(\mu)$  for different third order moments  $\mu_{(\bar{\mathbf{m}})}$  with  $\mathbf{R} = 2$  and  $n = 2$ .

$\mu_{\bar{\mathbf{m}}}$	$\varphi_{(\bar{\mathbf{m}})}(\mu)$
$\mathbf{E}[\mathbf{x}_1^3]$	$\left( \frac{\mathbf{E}[\mathbf{x}_1^2]}{\mathbf{E}[\mathbf{x}_1]} \right)^3$
$\mathbf{E}[\mathbf{x}_2^3]$	$\left( \frac{\mathbf{E}[\mathbf{x}_2^2]}{\mathbf{E}[\mathbf{x}_2]} \right)^3$
$\mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2]$	$\left( \frac{\mathbf{E}[\mathbf{x}_1^2]}{\mathbf{E}[\mathbf{x}_2]} \right) \left( \frac{\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2]}{\mathbf{E}[\mathbf{x}_1]} \right)^2$
$\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2^2]$	$\left( \frac{\mathbf{E}[\mathbf{x}_2^2]}{\mathbf{E}[\mathbf{x}_1]} \right) \left( \frac{\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2]}{\mathbf{E}[\mathbf{x}_2]} \right)^2$

$\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2^2]$  can be directly obtained from that of  $\mu_{(2,1)}$  by just switching the indices 1 and 2 in the expression for the moment closure function.

We now consider a third order truncation  $\mathbf{R} = 3$  for which

$$\begin{aligned}\boldsymbol{\mu} &= [\mu_{(1,0)}, \mu_{(0,1)}, \mu_{(2,0)}, \mu_{(0,2)}, \mu_{(1,1)}, \mu_{(3,0)}, \mu_{(0,3)}, \mu_{(2,1)}, \mu_{(1,2)}]^T \\ &:= [\mathbf{E}[\mathbf{x}_1], \mathbf{E}[\mathbf{x}_2], \mathbf{E}[\mathbf{x}_1^2], \mathbf{E}[\mathbf{x}_2^2], \mathbf{E}[\mathbf{x}_1 \mathbf{x}_2], \mathbf{E}[\mathbf{x}_1^3], \mathbf{E}[\mathbf{x}_2^3], \mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2], \mathbf{E}[\mathbf{x}_1 \mathbf{x}_2^2]]^T. \quad (1.47)\end{aligned}$$

Table 1.4 lists the different moment closure functions obtained for different fourth order moments of  $\mathbf{x}$  using the technique explained above. We only list a subset of the

Table 1.4. Moment closure function  $\varphi_{(\bar{\mathbf{m}})}(\boldsymbol{\mu})$  for different fourth order moments  $\mu_{(\bar{\mathbf{m}})}$  with  $\mathbf{R} = 3$  and  $n = 2$ .

$\mu_{\bar{\mathbf{m}}}$	$\varphi_{(\bar{\mathbf{m}})}(\boldsymbol{\mu})$
$\mathbf{E}[\mathbf{x}_1^3 \mathbf{x}_2]$	$\left( \frac{\mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2]}{\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2]} \right)^3 \left( \frac{\mathbf{E}[\mathbf{x}_1]}{\mathbf{E}[\mathbf{x}_1^2]} \right)^3 \mathbf{E}[\mathbf{x}_1^3] \mathbf{E}[\mathbf{x}_2]$
$\mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2^2]$	$\frac{(\mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2] \mathbf{E}[\mathbf{x}_1 \mathbf{x}_2^2] \mathbf{E}[\mathbf{x}_1] \mathbf{E}[\mathbf{x}_2])^2}{(\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2])^4 \mathbf{E}[\mathbf{x}_1^2] \mathbf{E}[\mathbf{x}_2^2]}$

moment closure functions. The remaining ones can be obtained from the ones shown by exchanging indices as discussed above.

### Three-species reactions ( $n = 3$ )

For a second order truncation ( $\mathbf{R} = 2$ ), the only third order moment whose moment closure function is not provided in Table 1.3 is  $\mu_{(1,1,1)} := \mathbf{E}[\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3]$ . Using Theorem 1.2 we obtain the following moment closure function in terms of the first and second order moments

$$\varphi_{(1,1,1)}(\boldsymbol{\mu}) = \frac{\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2] \mathbf{E}[\mathbf{x}_2 \mathbf{x}_3] \mathbf{E}[\mathbf{x}_1 \mathbf{x}_3]}{\mathbf{E}[\mathbf{x}_1] \mathbf{E}[\mathbf{x}_2] \mathbf{E}[\mathbf{x}_3]}. \quad (1.48)$$

Similarly, for a third order truncation ( $\mathbf{R} = 3$ ) we obtain the following moment closure function for  $\mu_{(2,1,1)} := \mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2 \mathbf{x}_3]$ :

$$\varphi_{(2,1,1)}(\mu) = \frac{(\mathbf{E}[\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3] \mathbf{E}[\mathbf{x}_1])^2 \mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_2] \mathbf{E}[\mathbf{x}_1^2 \mathbf{x}_3] \mathbf{E}[\mathbf{x}_2] \mathbf{E}[\mathbf{x}_3]}{(\mathbf{E}[\mathbf{x}_1 \mathbf{x}_3] \mathbf{E}[\mathbf{x}_1 \mathbf{x}_2])^2 \mathbf{E}[\mathbf{x}_1^2] \mathbf{E}[\mathbf{x}_2 \mathbf{x}_3]}. \quad (1.49)$$

One can see that all above moment closure functions are independent of the stoichiometry of the reactions, the reaction rates and the number of reactions. Another interesting observation is that the dependence of higher-order moment on lower order ones as given by all the above moment closure functions is consistent with  $\mathbf{x}$  being jointly *lognormally distributed*.

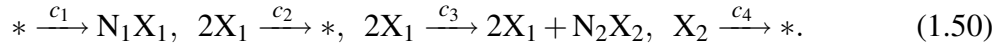
The procedure described here to generate approximated moment dynamics can be fully automated. The software `StochDynTools`, available at

[http : //www.ece.ucsb.edu/ ~ hespanha/software/stochdyntool.html](http://www.ece.ucsb.edu/~hespanha/software/stochdyntool.html)

computes truncated moment dynamics for any order of truncation starting from a simple ASCII description of the chemical reactions involved.

### 1.3 Example

In this section we consider the following set of chemical reactions



This example is motivated by a gene cascade network where a gene expresses a protein  $X_1$  and each expression event produces  $N_1$  molecules of the protein. The protein  $X_1$  then undergoes dimerization and activates another gene to express protein  $X_2$  with each expression event producing  $N_2$  molecules of the protein. The parameters given by  $c_1 = 100V$ ,  $c_2 = 0.8/V$ ,  $c_3 = 0.02/V$ ,  $c_4 = 15$  are defined in terms of the volume  $V$ ,

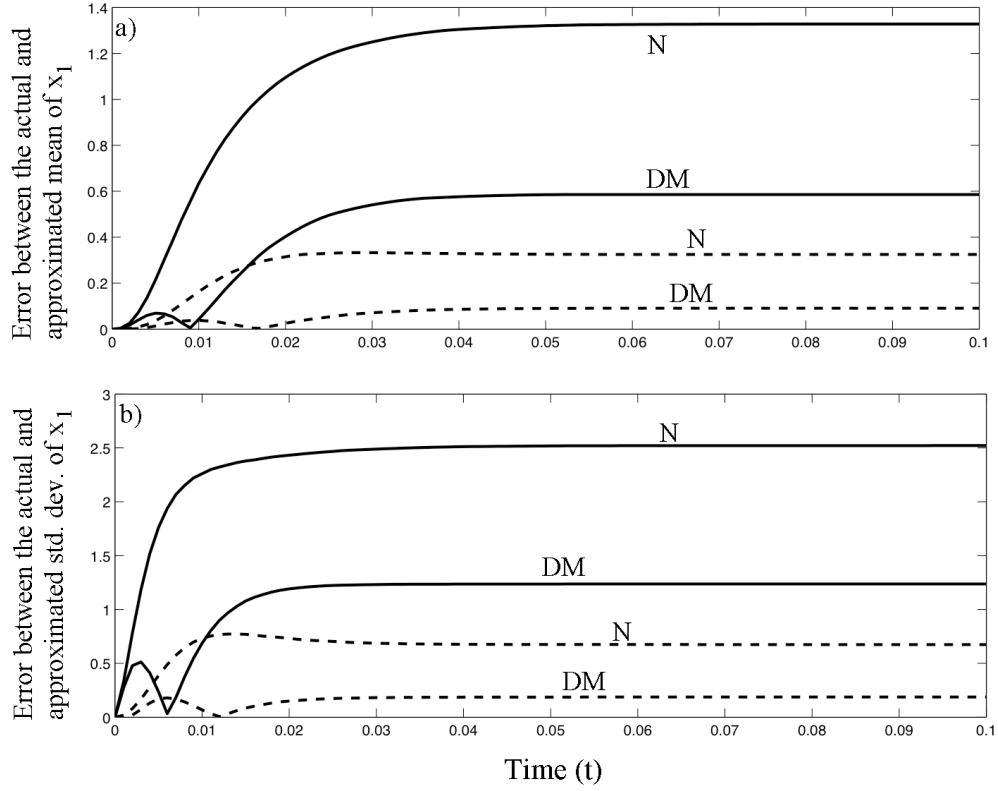


Figure 1.1. Plots of the absolute value of the error between the actual and approximated moment dynamics for a) the mean of  $\mathbf{x}_1$  and b) the standard deviation of  $\mathbf{x}_1$  corresponding to a second order truncation (solid line) and a third order truncation (dashed line) when the population size is small ( $V = 1$  units). DM and N, refer to the errors corresponding to the SDM moment closure and the normal moment closure, respectively. Other parameters taken as  $N_1 = 40$ ,  $N_2 = 15$  and initial conditions  $\mathbf{x}_1(0) = \mathbf{x}_2(0) = 40$ .

which is directly related to the population size of the chemical species. We consider two different moment closure methods: the SDM moment closure method developed in this chapter, and mass fluctuation kinetics where moment closure is performed by setting the third order *centered* moments equal to zero [16]. This later method is

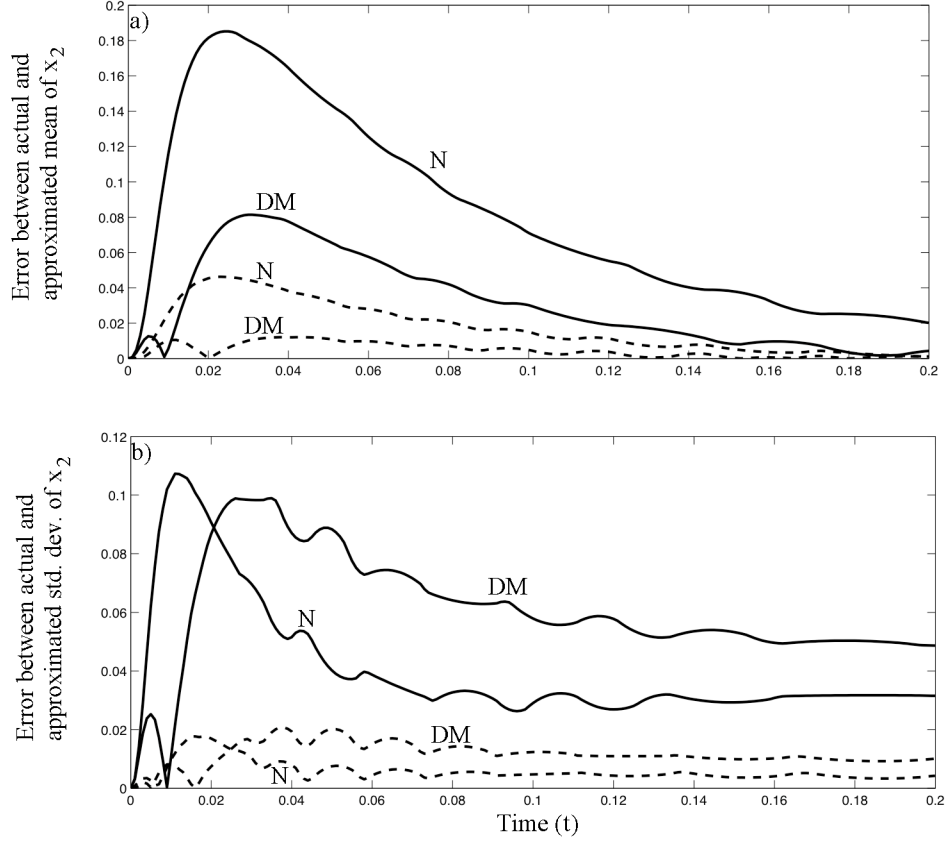


Figure 1.2. Plots of the absolute value of the error between the actual and approximated moment dynamics for a) the mean of  $\mathbf{x}_2$  and b) the standard deviation of  $\mathbf{x}_2$  corresponding to a second order truncation (solid line) and a third order truncation (dashed line) when the population size is small ( $V = 1$  units). DM and N, refer to the errors corresponding to the SDM moment closure and the normal moment closure, respectively. Other parameters taken as  $N_1 = 40$ ,  $N_2 = 15$  and initial conditions  $\mathbf{x}_1(0) = \mathbf{x}_2(0) = 40$ .

referred to as the *normal moment closure*. In particular, we see how these two different moment closure methods perform when the population size is small and large.

We first consider the situation where the population size is small and take  $V = 1$

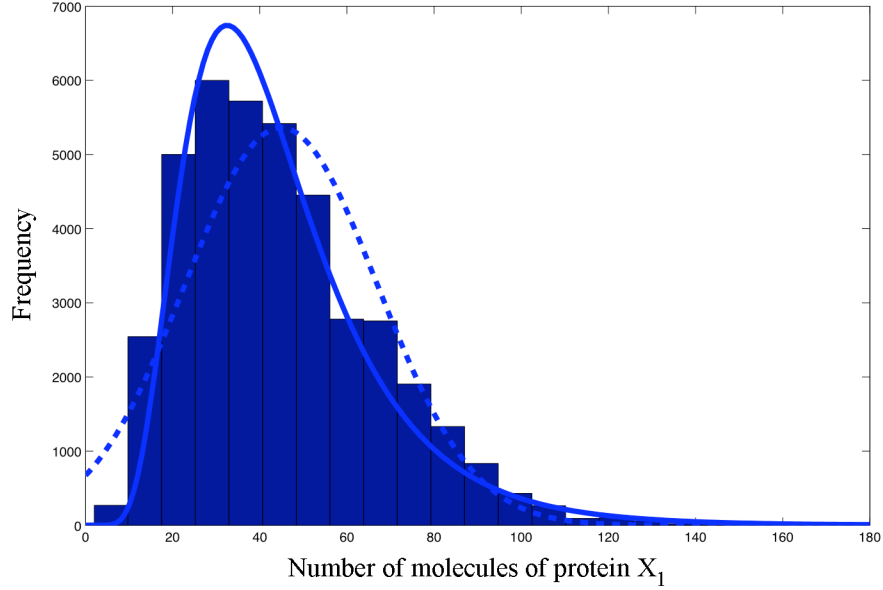


Figure 1.3. The steady-states histogram for  $\mathbf{x}_1$ , the number of molecules of protein  $X_1$ . The solid and dashed line corresponds to a lognormal and normal distribution approximation of the histogram, respectively, with mean and standard deviations obtained from the corresponding truncated moment dynamics

units which corresponds to a steady-state average number of  $\approx 50$  molecules, for both species. Figures 1.1 and 1.2 plot the moment estimates for species  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , respectively, corresponding to a second (solid lines) and a third (dashed lines) order truncation. As can be seen the SDM moment closure yields smaller errors between the approximated and actual moment dynamics compared to the normal moment closure. The explanation for this can be deduced from Figure 1.3, which plots the steady-state distribution of  $\mathbf{x}_1$ . This positively skewed distribution closely approximates a lognormal distribution, and hence, the SDM moment closure functions which are consistent with  $\mathbf{x}$  being jointly lognormally distributed provide better moment estimates. Also notice from Figure 1.3 that the corresponding normal distribution approximation (dashed

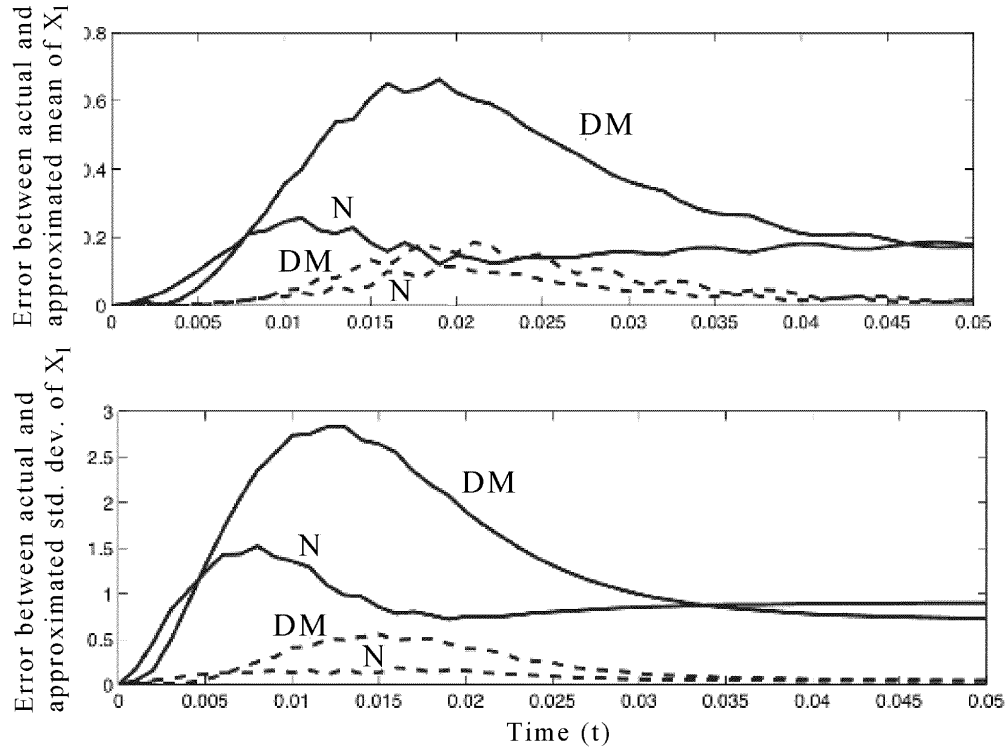


Figure 1.4. Plots of the absolute value of the error between the actual and approximated moment dynamics for a) the mean of  $\mathbf{x}_1$  and b) the standard deviation of  $\mathbf{x}_1$  corresponding to a second order truncation (solid line) and a third order truncation (dashed line) when the population size is large ( $V = 5$  units). DM and N, refer to the errors corresponding to the SDM moment closure and the normal moment closure, respectively. Other parameters taken as  $N_1 = 40$ ,  $N_2 = 15$  and initial conditions  $\mathbf{x}_1(0) = \mathbf{x}_2(0) = 40$ .

line) has some portion in the negative region, which is not biologically meaningful since molecule populations cannot drop below zero. Unlike the lognormal distribution which is only defined for positive values, this is frequently a problem with the normal approximation when the average number of molecules is small.

Figures 1.1 and 1.2 also confirm that increasing the order of truncation reduces the errors and considerably improves the moment estimates from the truncated moment dynamics. For example, with  $\mathbf{R} = 2$ , the error in the estimation of the steady-state standard deviation of  $\mathbf{x}_1$  is approximately 5.2% and 11% for the SDM and the normal moment closure method, respectively. However, with a third order truncation these errors reduce to 0.8% and 3.3%, respectively.

Finally, we consider the situation where the population size is large and increase the volume five-fold to  $V = 5$  units. Comparing Figure 1.4 with Figure 1.1 we see that in this case the normal moment closure method provides better moment estimates than the SDM moment closure method. This happens because when the population size is large the distribution is no longer positively skewed but symmetric about the mean. Indeed, moment closure functions consistent with a gaussian distribution perform better than those consistent with a lognormal distribution.

We refer interested readers to other examples that have been investigated using the SDM moment closure procedure. These include gene regulatory networks [51], stochastic models in population biology [50], Schögl reaction and other examples [19].

## 1.4 Conclusion

A procedure for estimating the time evolution of some lower order moments for the number of molecules of different species involved in a chemical reaction was presented. This was done by first explicitly writing the dynamics of these moments and then closing differential equations using moment closure. Moment closure was performed by first assuming a separable form for the moment closure function for each



element of  $\bar{\mu}$ , and then, matching it is time derivatives with  $\bar{\mu}$ , at some initial time  $t_0$  for a basis of initial conditions. The main result of this chapter is Theorem 1.2 which provides explicit formulae to compute the moment closure functions for higher order moments in the vector  $\bar{\mu}$  as a nonlinear function of the lower order moments in  $\mu$ , for any order of truncation  $\mathbf{R}$ . These moment closure functions are independent of the reaction parameters and the dependence of higher-order moment on lower order ones is consistent with the population being jointly *lognormally distributed*. These formulae can also be used to compute moment closure when the reactions in (1.4) are non-elementary (i.e., can have more than two reactants per reaction) although in such a case the degree of the polynomials in the error term  $\varepsilon_2(\mathbf{x}_0)$  would be larger than as stated in Theorem 1.2.

We constructed truncated moment dynamics for an example motivated by a gene cascade network. It provided fairly good estimates for the time evolution of the first and second order moments, with faster simulation times when compared to those needed to perform the same computations by averaging a large number of Monte Carlo simulations. Comparison between the SDM and normal moment closure showed key differences in their performance. In particular, when the number of molecules is large and the probability distribution of  $\mathbf{x}$  is symmetric, the normal moment closure provides better moments estimates than the SDM moment closure. However, as shown in Figure 1.3, for smaller population sizes the distribution is not symmetric but skewed toward zero. In this regime, SDM moment closure functions, which are consistent with a lognormal distribution, outperform the normal moment closure and yield smaller errors between the actual and approximated moment dynamics. Another drawback of the normal moment closure method is that in some cases, the corresponding truncated moment dynamics becomes dynamically unstable and blows up for sufficiently small

population sizes [32]. Finally, we should emphasize that the accuracy of these moment closure methods can be considerably improved by increasing the order of truncation  $\mathbf{R}$ . However, this comes at the cost of doing more computational analysis as the dimension of the truncated moment dynamics also increases with  $\mathbf{R}$ .

In summary, we developed a new moment closure procedure which is suited for reactions with small population sizes and non-gaussian probability distributions. These are ideal for bio-chemical processes within cells where mRNAs and protein occur in low levels. Moreover, many recent gene profiling studies have also revealed that the probability distribution of these species is not symmetric but positively skewed, and in many cases well approximated by a lognormal distribution [7]. This moment closure method will be a valuable tool in the hand of researches studying stochastic variability arising in the context of such bio-chemical reactions.

### 1.4.1 Future work

Possible directions of future work are as follows. It should be noted that the size of the vector  $\mu$  (denoted by  $k$ ) increases rapidly with the number of species  $n$ . Given  $n$  species we have  $\mathbf{C}_1^n$  first order moments,  $\mathbf{C}_1^n + \mathbf{C}_2^n$  second order moments,  $\mathbf{C}_1^n + 2\mathbf{C}_2^n + \mathbf{C}_3^n$  third order moments, and so on. Hence, if we consider a second ( $\mathbf{M} = \mathbf{2}$ ) or third ( $\mathbf{M} = \mathbf{3}$ ) order truncation, then  $k = n(3 + n)/2$  or  $k = n(11 + 6n + n^2)/6$ , respectively. A direction of future work is to investigate how one can perform model reduction using the fact the number the molecules of some species is very large, thus one can neglect the stochastic variations in that species. Also some reactions have fast dynamics and reach equilibrium quickly. In that case one could approximate the number of molecules of one specie in terms of the other and reduces the problem to  $n - 1$  species. This

technique is also referred to as the quasi steady-state approximation [43].

The error bound given by (1.26) tends to be very conservative and another direction of future work is to come up with new methods that will more accurately quantify the difference between the approximated and actual moment dynamics for a given moment closure method.

## Chapter 2

### Small noise approximation

In this chapter we introduce a new small noise approximation that provides analytical formulas relating the steady-state statistical moments to the parameters of the chemical reaction. These formulas not only predict the stochastic fluctuations about the mean but also the deviation of the mean from the solution to the chemical rate equation. Such approximate formulae are useful as they help develop a qualitative understanding of how noise level change in response to alterations of biologically meaningful parameters.

We demonstrate the application of this small noise approximation on two examples for which the steady-state average number of molecules is different from the steady-state solution of the chemical rate equation. In both the cases, the well known and widely used linear noise approximation (also called the Van Kampen's approximation and the Fluctuation dissipation theorem [59, 21]) fails in the sense that it predicts that the steady-state means are equal to the steady-state solution of the chemical rate equation. On the other hand, the small noise approximation accurately predicts the

steady-state means and higher order moments.

## 2.1 Theoretical formulation

We construct a vector  $\mu$  that contains all first and second order moments of  $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T$  (i.e., perform a second order truncation). Assuming that there are some chemical reactions with two or more reactants, the time derivative of  $\mu$  is given by

$$\dot{\mu} = \hat{\mathbf{a}} + A\mu + B\bar{\mu} \quad (2.1)$$

where  $\bar{\mu}$  is a vector containing moments of order three and higher. As described in the previous chapter, we close the above moment dynamics by approximating  $\bar{\mu}$  as a nonlinear function  $\bar{\varphi}$  of  $\mu$ . The steady-state moments  $\mu^*$  are then given as the solution to the following equation

$$\hat{\mathbf{a}} + A\mu^* + B\bar{\varphi}(\mu^*) = 0. \quad (2.2)$$

Since solving (2.2) for  $\mu^*$  in closed form is generally not possible, we use perturbation methods to compute the approximate steady-states. This is done by linearizing the left-hand-side of (2.2) about the steady-state solution to the chemical rate equation. Let  $\phi_{X_i}^*$  denote the steady-state average number of molecules of species  $X_i$  predicted from the chemical rate equation. We define

$$\mathbf{E}^*[\mathbf{x}_i] := \phi_{X_i}^* (1 + \varepsilon_{X_i}) \quad (2.3a)$$

$$\mathbf{E}^*[\mathbf{x}_i^2] := \mathbf{E}^*[\mathbf{x}_i]^2 (1 + CV_{X_i}^2) \quad (2.3b)$$

$$\mathbf{E}^*[\mathbf{x}_i \mathbf{x}_j] := \mathbf{E}^*[\mathbf{x}_i] \mathbf{E}^*[\mathbf{x}_j] \left( 1 + \kappa_{(X_i, X_j)} \right), \quad i \neq j \quad (2.3c)$$

where  $\mathbf{E}^*$  denotes the steady-state value of the respective moment and  $\varepsilon_{X_i}$ ,  $CV_{X_i}$ ,  $\kappa_{(X_i, X_j)}$  constants to be determined so as to make the above equation exact. Note that if the

steady state average number of molecules of  $\mathbf{x}_i$  was precisely equal to  $\phi_{X_i}^*$ , then  $\varepsilon_{X_i}$  would be zero; if the population of  $\mathbf{x}_i$  had zero variance, then the coefficient of variation  $CV_{X_i}$  would be zero; and if the populations  $\mathbf{x}_i, \mathbf{x}_j$  were not correlated then  $\kappa_{(X_i, X_j)}$  would be zero. Assuming  $\varepsilon_{X_i}$ ,  $CV_{X_i}^2$  and  $\kappa_{(X_i, X_j)}$  to be sufficiently small, we expand (2.3) using a Taylor series and ignore quadratic and higher order terms in  $\varepsilon_{X_i}$ ,  $CV_{X_i}^2$  and  $\kappa_{(X_i, X_j)}$ . Since (2.3a) is already linear in  $\varepsilon_{X_i}$ , we do not re-write this equation. For the remaining ones, we obtain

$$\mathbf{E}^*[\mathbf{x}_i^2] = \mathbf{E}^*[\mathbf{x}_i]^2 (1 + CV_{X_i}^2) \approx \phi_{X_i}^{*2} (1 + 2\varepsilon_{X_i} + CV_{X_i}^2) \quad (2.4a)$$

$$\mathbf{E}^*[\mathbf{x}_i \mathbf{x}_j] = \mathbf{E}^*[\mathbf{x}_i] \mathbf{E}^*[\mathbf{x}_j] (1 + \kappa_{(X_i, X_j)}) \approx \phi_{X_i}^* \phi_{X_j}^* (1 + \varepsilon_{X_i} + \varepsilon_{X_j} + \kappa_{(X_i, X_j)}). \quad (2.4b)$$

Using the moment closure functions derived in the previous section we now express moments of order three and higher as linear combinations of  $\varepsilon_{X_i}$ ,  $CV_{X_i}^2$  and  $\kappa_{(X_i, X_j)}$ . We recall from Table 1.3 that the moment closure functions for the third order moments  $\mathbf{E}[\mathbf{x}_i^3]$  and  $\mathbf{E}[\mathbf{x}_i^2 \mathbf{x}_j]$  are given by

$$\mathbf{E}[\mathbf{x}_i^3] \approx \left( \frac{\mathbf{E}[\mathbf{x}_i^2]}{\mathbf{E}[\mathbf{x}_i]} \right)^3 \quad (2.5a)$$

$$\mathbf{E}[\mathbf{x}_i^2 \mathbf{x}_j] \approx \frac{\mathbf{E}[\mathbf{x}_i^2]}{\mathbf{E}[\mathbf{x}_j]} \left( \frac{\mathbf{E}[\mathbf{x}_i \mathbf{x}_j]}{\mathbf{E}[\mathbf{x}_i]} \right)^2. \quad (2.5b)$$

Using (2.3) and (2.4) we have that

$$\mathbf{E}[\mathbf{x}_i^3] \approx \left( \frac{\mathbf{E}[\mathbf{x}_i^2]}{\mathbf{E}[\mathbf{x}_i]} \right)^3 \approx \phi_{X_i}^{*3} (1 + 3\varepsilon_{X_i} + 3CV_{X_i}^2) \quad (2.6a)$$

$$\mathbf{E}[\mathbf{x}_i^2 \mathbf{x}_j] \approx \frac{\mathbf{E}[\mathbf{x}_i^2]}{\mathbf{E}[\mathbf{x}_j]} \left( \frac{\mathbf{E}[\mathbf{x}_i \mathbf{x}_j]}{\mathbf{E}[\mathbf{x}_i]} \right)^2 \approx \phi_{X_i}^{*2} \phi_{X_j}^* (1 + 2\varepsilon_{X_i} + \varepsilon_{X_j} + CV_{X_i}^2 + 2\kappa_{(X_i, X_j)}). \quad (2.6b)$$

In general, using Theorem 1.2 one can show that any higher order moment can be written as a linear combinations of  $\varepsilon_{X_i}$ ,  $CV_{X_i}^2$ ,  $\kappa_{(X_i, X_j)}$  as follows:

$$\begin{aligned} \mu_{(\mathbf{m})} &:= \mathbf{E}[\mathbf{x}_1^{m_1} \dots \mathbf{x}_n^{m_n}] \approx \varphi_{(\mathbf{m})}(\mu) \approx \\ &\phi_{X_1}^{* m_1} \dots \phi_{X_n}^{* m_n} \left( 1 + \sum_{j=1}^n m_j \varepsilon_{X_j} + \sum_{j=1}^n \frac{m_j(m_j-1)}{2} CV_{X_j}^2 + \sum_{j=1}^n \sum_{k=1(j \neq k)}^n m_j m_k \kappa_{(X_i, X_j)} \right) \end{aligned} \quad (2.7)$$

where  $\mathbf{m} = (m_1, \dots, m_n)$ . Substituting (2.7) in (2.2) we obtain a linear system of equations which can be analytically solved to obtain the steady-state moments. In the remaining of this chapter we illustrate the use of this small noise approximation on two different examples and compare it with the widely used linear noise approximation.

## 2.2 Gene expression and protein degradation

We consider the following set of chemical reactions



These reactions are motivated by a gene expression process in which the protein  $X$  is expressed from the gene at a constant rate  $c_1$ . Each gene expression event leads to the formation of  $N_x$  molecules of  $X$ .  $D$  molecules of the protein then combine to form a multimer which degrades at a constant rate  $c_2$ . As before, we model the number of molecules  $\mathbf{x}$  of the protein by a SHS with trivial dynamics and two reset maps

$$\mathbf{x} \mapsto \phi_1(\mathbf{x}) = \mathbf{x} + N_x, \quad \mathbf{x} \mapsto \phi_2(\mathbf{x}) = \mathbf{x} - D \quad (2.9)$$

with corresponding transition intensities given by

$$\lambda_1(\mathbf{x}) = c_1, \quad \lambda_2(\mathbf{x}) = c_2 \mathbf{x}^D. \quad (2.10)$$

The solution  $\phi_X$  of the corresponding chemical rate equation is given by

$$\frac{d\phi_X}{dt} = c_1 N_x - c_2 D \phi_X^D \quad (2.11)$$

which predicts the following steady-state average number of protein molecules

$$\phi_X^* = \left( \frac{c_1 N_x}{D c_2} \right)^{1/D}. \quad (2.12)$$

### 2.2.1 Small noise approximation

The exact time derivative of the first and second order moment of  $\mathbf{x}$  is given by

$$\frac{d\mathbf{E}[\mathbf{x}]}{dt} = c_1 N_x - c_2 D \mathbf{E}[\mathbf{x}^D] \quad (2.13a)$$

$$\frac{d\mathbf{E}[\mathbf{x}^2]}{dt} = c_1 N_x^2 + 2c_1 N_x \mathbf{E}[\mathbf{x}] + c_2 D^2 \mathbf{E}[\mathbf{x}^D] - 2c_2 D \mathbf{E}[\mathbf{x}^{D+1}] \quad (2.13b)$$

and therefore the steady-state moments are the solution to

$$0 = c_1 N_x - c_2 D \mathbf{E}^*[\mathbf{x}^D] \quad (2.14a)$$

$$0 = c_1 N_x^2 + 2c_1 N_x \mathbf{E}^*[\mathbf{x}] + c_2 D^2 \mathbf{E}^*[\mathbf{x}^D] - 2c_2 D \mathbf{E}^*[\mathbf{x}^{D+1}] \quad (2.14b)$$

where  $\mathbf{E}^*$  denotes the steady-state value of the respective moment. We have from (2.14a) and Jensen's inequality that

$$\frac{c_1 N_x}{c_2 D} = \mathbf{E}^*[\mathbf{x}^D] > \mathbf{E}^*[\mathbf{x}]^D \quad (2.15)$$

which implies from (2.12)

$$\mathbf{E}^*[\mathbf{x}] < \left( \frac{c_1 N_x}{D c_2} \right)^{1/D} = \phi_X^*. \quad (2.16)$$

With this we conclude that the chemical rate equation over-estimates the steady-state value for the average number of protein molecules.



Defining

$$\mathbf{E}^*[\mathbf{x}] := \phi_X^* (1 + \varepsilon_X) \quad (2.17a)$$

$$\mathbf{E}^*[\mathbf{x}^2] := \mathbf{E}^*[\mathbf{x}]^2 (1 + CV_X^2) \quad (2.17b)$$

and assuming  $\varepsilon_X^2 \ll 1$ ,  $CV_X^4 \ll 1$ ,  $\varepsilon_X CV_X^2 \ll 1$ , we have from (2.7) that

$$\mathbf{E}^*[\mathbf{x}^D] \approx \phi_X^{*D} \left( 1 + D\varepsilon_X + \frac{D(D-1)}{2} CV_X^2 \right) \quad (2.18a)$$

$$\mathbf{E}^*[\mathbf{x}^{D+1}] \approx \phi_X^{*D+1} \left( 1 + (D+1)\varepsilon_X + \frac{D(D+1)}{2} CV_X^2 \right). \quad (2.18b)$$

Substituting the above moment approximations in (2.14), we obtain a linear system of equations in  $\varepsilon_X$  and  $CV_X^2$ . Analytically solving these equations yield the following steady-state moments

$$\mathbf{E}^*[\mathbf{x}] = \phi_X^* (1 + \varepsilon_X), \quad \varepsilon_X \approx -\frac{(D-1)(D+N_X)}{4D\phi_X^*} \quad (2.19a)$$

$$\mathbf{E}^*[\mathbf{x}^2] = \mathbf{E}^*[\mathbf{x}]^2 (1 + CV_X^2), \quad CV_X^2 \approx \frac{D+N_X}{2D\phi_X^*}. \quad (2.19b)$$

Note that our small noise approximation successfully predicts that the average number of molecules  $\mathbf{E}^*[\mathbf{x}]$  is smaller than the steady-solution  $\phi_X^*$  of the chemical rate equation.

### 2.2.2 Comparison with linear noise approximation

For comparison purposes, we perform the linear noise approximation by first linearizing the propensity functions about the solution of the chemical rate equation [59].

The transitional intensities (2.10) are now modified as

$$\lambda_1(\mathbf{x}) = c_1, \quad \lambda_2(\mathbf{x}) = c_2 (\phi_X^D + D\phi_X^{D-1}(\mathbf{x} - \phi_X)) \quad (2.20)$$

where  $\phi_X$  is the solution to (2.11). Writing the moment dynamics for these linear transitional intensities results in the following closed system of differential equations

$$\frac{d\mathbf{E}[\mathbf{x}]}{dt} = c_1 N_x - c_2 D \phi_X^D + D \phi_X^{D-1} (\mathbf{E}[\mathbf{x}] - \phi_X) \quad (2.21a)$$

$$\begin{aligned} \frac{d\mathbf{E}[\mathbf{x}^2]}{dt} = & c_1 N_x^2 + 2c_1 N_x \mathbf{E}[\mathbf{x}] + c_2 \phi_X^D (D^2 - 2D\mathbf{E}[\mathbf{x}]) \\ & + c_2 D^2 \phi_X^{D-1} (D\mathbf{E}[\mathbf{x}] + 2\phi_X \mathbf{E}[\mathbf{x}] - D\phi_X - 2\mathbf{E}[\mathbf{x}^2]). \end{aligned} \quad (2.21b)$$

Note from (2.21a) that the dynamics of the mean is now independent of  $\mathbf{E}[\mathbf{x}^2]$ . A steady-state analysis of (2.21) gives the following steady-state moments

$$\mathbf{E}^*[\mathbf{x}] \approx \phi_X^* \quad (2.22a)$$

$$\mathbf{E}^*[\mathbf{x}^2] = \mathbf{E}^*[\mathbf{x}]^2 (1 + CV_X^2), \quad CV_X^2 \approx \frac{D + N_X}{2D\phi_X^*}. \quad (2.22b)$$

Comparing (2.19) with (2.22) we see that both the approximations provide the same formula for the steady-state coefficient of variation. However, unlike the small noise approximation, (2.22a) erroneously predicts that the average number of molecules  $\mathbf{E}^*[\mathbf{x}]$  is equal to the steady-solution  $\phi_X^*$  of the chemical rate equation.

Monte Carlo simulations confirm that the small noise approximation provides more accurate estimates of  $\mathbf{E}^*[\mathbf{x}]$  than the linear noise approximation (see Table 2.1). Figure 2.1 shows that a lognormal distribution with moment estimates from the small noise approximation provides a very good match to the actual steady-state distribution of the protein population.

## 2.3 Gene expression and activation

We now consider the following set of chemical reactions



Table 2.1. Steady-state mean ( $\mathbf{E}^*[\mathbf{x}]$ ) and coefficient of variation ( $CV_X$ ) from the small noise approximation (SNA) presented in this chapter, linear noise approximation (LNA) and Monte Carlo simulations (MC). The following parameters were used:  $N_x = 35$ ,  $D = 5$  and  $\phi_X^* = 50$ .

	SNA	LNA	MC
$\mathbf{E}^*[\mathbf{x}]$	42	50	$\approx 42.7$
$CV_X$	0.283	0.283	$\approx 0.288$

which corresponds to a gene expressing a protein  $X_1$ .  $D \geq 2$  molecules of the protein then combine to form a multimer that activates another gene to make protein  $X_2$ . The above reactions can be modelled by the following reset maps

$$\mathbf{x} \mapsto \phi_1(\mathbf{x}) = \begin{bmatrix} \mathbf{x}_1 + 1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{x} \mapsto \phi_2(\mathbf{x}) = \begin{bmatrix} \mathbf{x}_1 - 1 \\ \mathbf{x}_2 \end{bmatrix}, \quad (2.24a)$$

$$\mathbf{x} \mapsto \phi_3(\mathbf{x}) = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 + 1 \end{bmatrix}, \quad \mathbf{x} \mapsto \phi_4(\mathbf{x}) = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 - 1 \end{bmatrix}, \quad (2.24b)$$

with the following transition intensities

$$\lambda_1(\mathbf{x}) = c_1, \quad \lambda_2(\mathbf{x}) = c_2 \mathbf{x}_1, \quad \lambda_3(\mathbf{x}) = c_3 \mathbf{x}_1^D, \quad \lambda_4(\mathbf{x}) = c_4 \mathbf{x}_2. \quad (2.25)$$

The corresponding chemical rate equations are given by

$$\frac{d\phi_{X_1}}{dt} = c_1 - c_2 \phi_{X_1} \quad (2.26a)$$

$$\frac{d\phi_{X_2}}{dt} = c_3 \phi_{X_1}^D - c_4 \phi_{X_2} \quad (2.26b)$$

where  $\phi_{X_1}$  and  $\phi_{X_2}$  are the predicted number of molecules of protein  $X_1$  and  $X_2$ , respectively, from the deterministic model. Steady-state analysis of the above equations

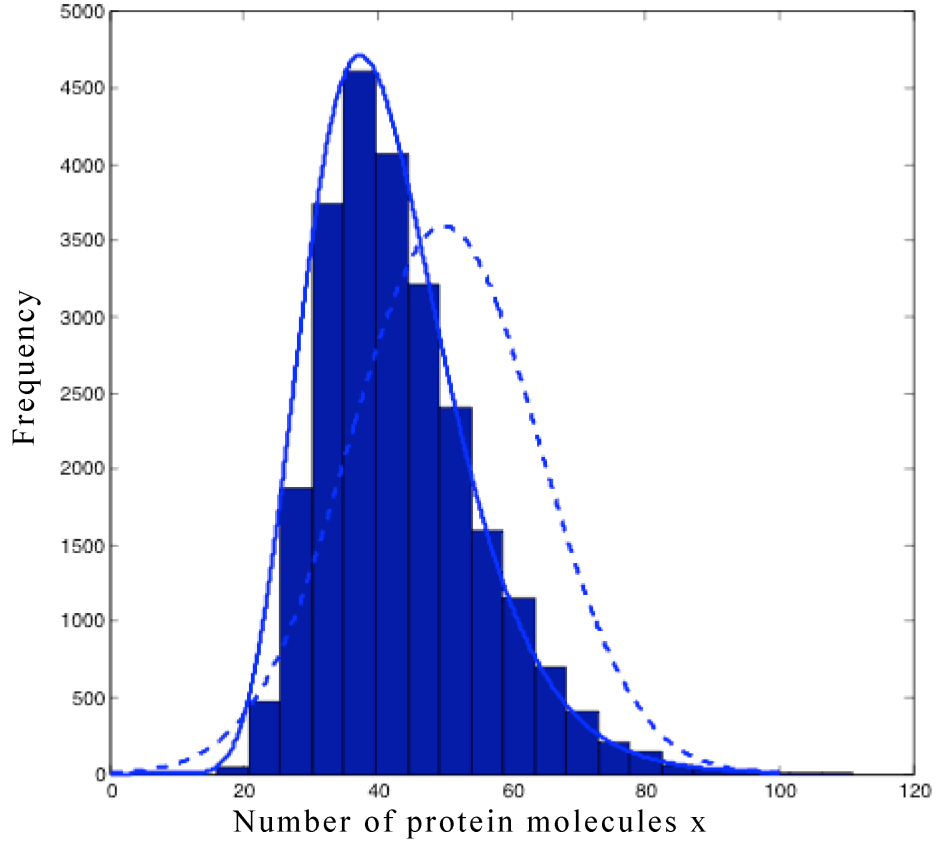


Figure 2.1. Steady-state histogram of  $\mathbf{x}$  obtained from 10,000 Monte Carlo simulations. Solid line is a lognormal approximation using the steady-state moments from the small noise approximation (equation (2.19)). Dashed line is a normal approximation using the steady-state moments from the linear noise approximation (equation (2.22))

yields

$$\phi_{X_1}^* = \frac{c_1}{c_2}, \quad \phi_{X_2}^* = \frac{c_3 \phi_{X_1}^{*D}}{c_4}. \quad (2.27)$$

The time evolution of the means  $\mathbf{E}[\mathbf{x}_1]$  and  $\mathbf{E}[\mathbf{x}_2]$  are given by

$$\frac{d\mathbf{E}[\mathbf{x}_1]}{dt} = c_1 - c_2 \mathbf{E}[\mathbf{x}_1] \quad (2.28a)$$

$$\frac{d\mathbf{E}[\mathbf{x}_2]}{dt} = c_3 \mathbf{E}[\mathbf{x}_1^D] - c_4 \mathbf{E}[\mathbf{x}_2] \quad (2.28b)$$

which show that

$$\mathbf{E}^*[\mathbf{x}_1] = \phi_{X_1}^* \quad (2.29)$$

and from Jensen's inequality

$$\mathbf{E}^*[\mathbf{x}_2] = \frac{c_3}{c_4} \mathbf{E}^*[\mathbf{x}_1^D] > \frac{c_3}{c_4} \mathbf{E}^*[\mathbf{x}_1]^D = \phi_{X_2}^*. \quad (2.30)$$

We conclude from above that the steady-state average number of molecules of species  $X_2$  is larger than its corresponding deterministic approximation  $\phi_{X_2}^*$ . We next use small noise approximation to predict the difference between  $\mathbf{E}^*[\mathbf{x}_2]$  and  $\phi_{X_2}^*$ .

### 2.3.1 Small noise approximation

The time derivative of all the second order moments of  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2]^T$  are given by

$$\frac{d\mathbf{E}[\mathbf{x}_1^2]}{dt} = c_1 + 2c_1\mathbf{E}[\mathbf{x}_1] + c_2\mathbf{E}[\mathbf{x}_1] - 2c_2\mathbf{E}[\mathbf{x}_1^2] \quad (2.31a)$$

$$\frac{d\mathbf{E}[\mathbf{x}_1\mathbf{x}_2]}{dt} = c_3\mathbf{E}[\mathbf{x}_1^{D+1}] + c_1\mathbf{E}[\mathbf{x}_2] - c_2\mathbf{E}[\mathbf{x}_1\mathbf{x}_2] - c_4\mathbf{E}[\mathbf{x}_1\mathbf{x}_2] \quad (2.31b)$$

$$\frac{d\mathbf{E}[\mathbf{x}_2^2]}{dt} = c_3\mathbf{E}[\mathbf{x}_1^D] + c_4\mathbf{E}[\mathbf{x}_2] + 2c_4\mathbf{E}[\mathbf{x}_1^D\mathbf{x}_2] - 2c_4\mathbf{E}[\mathbf{x}_2^2]. \quad (2.31c)$$

We solve the steady-state moments by setting the left-hand-side of equations (2.28) and (2.31) to zero and from (2.7) substituting

$$\mathbf{E}^*[\mathbf{x}_1^D] \approx \phi_{X_1}^{*D} \left( 1 + D\varepsilon_{X_1} + \frac{D(D-1)}{2} CV_{X_1}^2 \right) \quad (2.32a)$$

$$\mathbf{E}^*[\mathbf{x}_1^{D+1}] \approx \phi_{X_1}^{*D+1} \left( 1 + (D+1)\varepsilon_{X_1} + \frac{D(D+1)}{2} CV_{X_1}^2 \right) \quad (2.32b)$$

$$\mathbf{E}^*[\mathbf{x}_1^D\mathbf{x}_2] \approx \phi_{X_1}^{*D}\phi_{X_2}^* \left( 1 + D\varepsilon_{X_1} + \varepsilon_{X_2} + \frac{D(D-1)}{2} CV_{X_1}^2 + D\kappa_{(X_1, X_2)} \right). \quad (2.32c)$$

Solving the resulting linear system of equations gives the following steady-state moments

$$\mathbf{E}^*[\mathbf{x}_1] = \phi_{X_1}^*, \quad CV_{X_1}^2 = \frac{1}{\phi_{X_1}^*} \quad (2.33a)$$

$$\mathbf{E}^*[\mathbf{x}_2] = \phi_{X_2}^* (1 + \varepsilon_{X_2}), \quad \varepsilon_{X_2} \approx \frac{D(D-1)}{2\phi_{X_1}^*}, \quad (2.33b)$$

$$CV_{X_2}^2 \approx \frac{1}{\phi_{X_2}^*} + \frac{D^2}{\phi_{X_1}^* + \phi_{X_2}^*} + \frac{D(D-1)}{2\phi_{X_2}^* \phi_{X_1}^*} \quad (2.33c)$$

$$\mathbf{E}^*[\mathbf{x}_1 \mathbf{x}_2] = \mathbf{E}^*[\mathbf{x}_1] \mathbf{E}^*[\mathbf{x}_2] (1 + \kappa_{(X_1, X_2)}), \quad \kappa_{(X_1, X_2)} \approx \frac{D}{\phi_{X_1}^* + \phi_{X_2}^*}. \quad (2.33d)$$

Note from (2.33b) that the steady-state average number of molecules of the protein  $X_2$  is larger than  $\phi_{X_2}^*$  by  $\phi_{X_2}^* D(D-1)/2\phi_{X_1}^*$ . This phenomenon, where noise causes the average signal to be larger than its corresponding deterministic signal is often referred to as *stochastic focusing* [39].

### 2.3.2 Comparison with linear noise approximation

We perform linear noise approximation by linearizing the propensity functions (2.25) about the solution to the chemical rate equations given by (2.26). The linearized propensity functions are given by

$$\lambda_1(\mathbf{x}) = c_1, \quad \lambda_2(\mathbf{x}) = c_2 \mathbf{x}_1, \quad \lambda_3(\mathbf{x}) = c_3 \left( \phi_{X_1}^D + D \phi_{X_1}^{D-1} (\mathbf{x}_1 - \phi_{X_1}^*) \right), \quad \lambda_4(\mathbf{x}) = c_4 \mathbf{x}_2. \quad (2.34)$$

Moment dynamics with these linearized propensity functions yields the following steady-state moments

$$\mathbf{E}^*[\mathbf{x}_1] = \phi_{X_1}^*, \quad CV_{X_1}^2 = \frac{1}{\phi_{X_1}^*} \quad (2.35a)$$

$$\mathbf{E}^*[\mathbf{x}_2] = \phi_{X_2}^*, \quad CV_{X_2}^2 \approx \frac{1}{\phi_{X_2}^*} + \frac{D^2}{\phi_{X_1}^* + \phi_{X_2}^*} \quad (2.35b)$$

$$\mathbf{E}^*[\mathbf{x}_1 \mathbf{x}_2] = \mathbf{E}^*[\mathbf{x}_1] \mathbf{E}^*[\mathbf{x}_2] (1 + \kappa_{(X_1, X_2)}) , \quad \kappa_{(X_1, X_2)} \approx \frac{D}{\phi_{X_1}^* + \phi_{X_2}^*}. \quad (2.35c)$$

As expected, the linear noise approximation erroneously predicts the difference between  $\mathbf{E}^*[\mathbf{x}_2]$  and its deterministic approximation  $\phi_{X_2}^*$  to be zero. However, opposite to what happened in the previous example, now the linear noise approximation yields a different formula for  $CV_{X_2}$  than our small noise approximation:

$$CV_{X_2}^2 \approx \frac{1}{\phi_{X_2}^*} + \frac{D^2}{\phi_{X_1}^* + \phi_{X_2}^*} + \frac{D(D-1)}{2\phi_{X_2}^* \phi_{X_1}^*} \quad (\text{Small noise approximation}) \quad (2.36a)$$

$$CV_{X_2}^2 \approx \frac{1}{\phi_{X_2}^*} + \frac{D^2}{\phi_{X_1}^* + \phi_{X_2}^*} \quad (\text{Linear noise approximation}), \quad (2.36b)$$

however, the difference between the two predicted coefficient of variations is very small for  $\phi_{X_1}^*, \phi_{X_2}^* \gg 1$ .

## 2.4 Conclusion

This chapter introduced a new small noise approximation which provides analytical formulas for the steady-state moments in terms of the reaction parameters. These formulas are more accurate than those obtained with the well known linear noise approximation as they take into account the effect of the noise in the dynamics of the mean. We demonstrated the small noise approximation on two examples, where it successfully predicted not only the fluctuations about the mean but also the deviation of the mean from the solution of the chemical rate equation.

## Chapter 3

# Moment closure for the stochastic logistic model

In chapter 1 we developed a new moment closure procedure that provides the time evolution of all the lower-order statistical moments of the species population. We now use this procedure to estimate the statistical moments for a special class of birth-death process known as the *stochastic logistic model*. This model has been widely used in ecology for stochastic modeling of single-species population in environments with constrained resources [26, 34]. Details of the stochastic logistic model are presented in Section 3.1.

It is well known that for the stochastic logistic model  $\mathbf{x} = 0$  is an absorbing state and eventual convergence to the origin is certain. For most biologically relevant problems one is typically interested in the distribution of the process conditioned on the event that absorption has not yet occurred. Let  $P_x(t) = \Pr\{\mathbf{x}(t) = x \mid \mathbf{x}(t) > 0\}$  denote the probability density function of the conditioned processes and  $\mu_m(t) = \sum_{x=1}^{\infty} x^m P_x(t)$  its



$m^{th}$  order uncentered moment. We show in Section 3.2 that the time derivative of the vector  $\mu = [\mu_1, \dots, \mu_k]^T \in \mathbb{R}^k$ , where  $k$  is the order of the truncation, is given by

$$\dot{\mu} = (A + \lambda_{\text{ext}}(t)I)\mu + B\mu_{k+1} \quad (3.1)$$

where  $A$  and  $B$  are appropriately defined matrices,  $I$  is the identity matrix, and  $\lambda_{\text{ext}}(t)$  an extinction rate. Assuming that the mean time to extinction is very large, the perturbation term  $\lambda_{\text{ext}}(t)I$  is very small when compared to the matrix  $A$  and can be ignored [32]. In spite of this, the dynamics of the above system is not closed, in the sense that the time evolution of the vector  $\mu$  depends on the  $k+1^{th}$  order moment  $\mu_{k+1}$ . We recall that the above moment dynamics is closed by performing moment closure which involves approximating  $\mu_{k+1}$  as a nonlinear function  $\varphi_{k+1}(\mu)$  of the moments up to order  $k$ . The resulting closed moment dynamics is then given by

$$\dot{\mathbf{v}} = A\mathbf{v} + B\varphi_{k+1}(\mathbf{v}). \quad (3.2)$$

In Section 3.3, we perform a Separable Derivative-Matching (SDM) moment closure. As illustrated before, this involves choosing a moment closure functions which has a separable form:

$$\varphi_{k+1}^s(\mathbf{v}) = v_1^{\gamma_1} v_2^{\gamma_2} \dots v_k^{\gamma_k} \quad (3.3)$$

for appropriately chosen constants  $\gamma_p \in \mathbb{R}$ . These constants are then obtained by matching time derivatives of  $\mu_{k+1}$  and  $\varphi_{k+1}^s(\mathbf{v})$  in (3.1) and (3.2) respectively, at some initial time  $t_0$  and initial condition  $\mathbf{x}(t_0) = x_0$  with probability one. In this chapter we denote the SDM moment closure functions with the superscript  $s$ , in order to distinguish them with other moment closure functions to be introduced later.

Alternative moment closure methods that have appeared in literature typically construct the moment closure functions  $\varphi$  by directly assuming the probability distribution

to be normal [64, 33], lognormal [22], Poisson or binomial [32]. We refer to them as *normal*, *lognormal*, *Poisson* and *binomial moment closures* respectively and review them in Section 3.4. In Section 3.5, they are compared with the SDM moment closure based on how well the moment closure function  $\phi_{k+1}(\mu)$  approximates  $\mu_{k+1}$ . Towards that end, we introduce the error

$$e_{k+1}(t) := \mu_{k+1}(t) - \phi_{k+1}(\mu(t)) = \sum_{i=0}^{\infty} \frac{(t-t_0)^i}{i!} \varepsilon_{k+1}^i(x_0) \quad (3.4)$$

where we expanded the error as a Taylor series with  $\varepsilon_{k+1}^i(x_0)$  defined to be

$$\varepsilon_{k+1}^i(x_0) := \left. \frac{d^i \mu_{k+1}(t)}{dt^i} \right|_{t=t_0} - \left. \frac{d^i \phi_{k+1}(\mu(t))}{dt^i} \right|_{t=t_0}. \quad (3.5)$$

We call  $\varepsilon_{k+1}^i(x_0)$  the  $i^{\text{th}}$  order *derivative matching error*. Ideally one would like this error to be zero but this is generally not possible. When  $\mathbf{x}(t_0) = x_0$  with probability one, the derivative matching error is typically a polynomial in  $x_0$ . For example, for the SDM moment closure function the  $0^{\text{th}}$  order derivative matching error is zero while for  $i \geq 1$  the  $i^{\text{th}}$  order error is a polynomial in  $x_0$  of degree  $i+1$ . Typically, the lesser the order of this polynomial, the lesser is the error  $e_{k+1}(t)$ , and hence the better is  $\phi_{k+1}(\mu)$  in approximating  $\mu_{k+1}$ .

We show that for  $k = 2$ , all the above moment closure functions perform derivative matching except the Poisson moment closure function proposed by [32]. This is because, it has a  $0^{\text{th}}$  order derivative matching error  $\varepsilon_3^0(x_0)$  which grows linearly with  $x_0$  while for SDM, lognormal, binomial and normal moment closure functions the  $0^{\text{th}}$  order error is always zero. Hence, the Poisson moment closure function proposed by [32] exhibits a larger initial error than the others. We propose an alternative Poisson moment closure function, for which  $\varepsilon_3^0(x_0) = 0$ , and show that it performs better than the one proposed by [32].

Although the above moment closures provide good estimates for a second order

of truncation ( $k = 2$ ), it is typically beneficial to consider higher order of truncations because they lead to better moment approximations and reduce the errors by a few orders of magnitude. To the authors knowledge, moment closure for  $k \geq 3$  has always been done in literature by assuming a normal distribution for the population [33, 25]. We show that for  $k = 3$  the normal moment closure function also performs derivative matching with similar derivative matching error as for the SDM moment closure function, and hence, gives fairly good estimates of  $\mu_4$ . However, for  $k = 3$  we further propose a new moment closure function that yields lesser derivative matching errors when compared to separable derivative matching and normal moment closure functions, thus providing better estimates for  $\mu_4$ , at least locally in time.

In Section 3.6 we find the steady-state solutions of the truncated moment dynamics (3.2)–(3.3). We show that the separable structure of the SDM moment closure leads to analytical expressions for the approximate steady-state moments  $v^*$ , which are always unique, real and positive for every order of truncation  $k \geq 2$ . In contrast, finding expressions for the steady-state moments using normal moment closure is typically done numerically for  $k > 2$ , as this involves solving an  $k^{th}$  degree polynomial and then identifying the biologically relevant steady-state among the  $k$  roots of the polynomial [25].

## 3.1 Stochastic logistic model

### 3.1.1 Model formulation

The stochastic logistic model is the stochastic birth-death analogous model of the well-known deterministic Verhulst-Pearl equations [40, 41] and has been exten-

sively used for modeling stochasticity in population biology [26, 25, 27, 23]. For this continuous-time birth-death Markov process, the conditional probabilities of a unit increase and decrease, respectively, in an “infinitesimal” time interval  $(t, t + dt]$  are given by

$$\Pr\{\mathbf{x}(t + dt) = x + 1 \mid \mathbf{x}(t) = x\} = \begin{cases} \eta(x)dt, & \forall x \leq U \\ 0, & \text{otherwise} \end{cases} \quad (3.6a)$$

$$\Pr\{\mathbf{x}(t + dt) = x - 1 \mid \mathbf{x}(t) = x\} = \chi(x)dt, \quad (3.6b)$$

where  $\mathbf{x}(t)$  represents the population size at time  $t$ ,

$$\eta(x) := a_1x - b_1x^2 > 0, \quad \chi(x) := a_2x + b_2x^2 > 0, \quad \forall x \in (0, U) \quad (3.7)$$

and

$$U := a_1/b_1, \quad a_1 > 0, \quad a_2 > 0, \quad b_1 > 0, \quad b_2 \geq 0. \quad (3.8)$$

We assume that the initial condition satisfies  $\mathbf{x}(t_0) \in \{1, 2, \dots, U\}$ , and hence,  $\mathbf{x}(t) \in \{0, 1, \dots, U\}, \forall t \in [0, \infty)$  with probability one. We call  $U$  the *population limit*.

### 3.1.2 Stationary and quasi-stationary distributions

Since the birth and death rates are zero for  $x = 0$  ( $\eta(0) = \chi(0) = 0$ ) we have that the state  $\mathbf{x} = 0$  is absorbing and eventual convergence to the origin is certain. However, it is common to use the stochastic logistic model mainly in the case where the mean time to extinction is very large. As the stationary distribution is degenerate with probability one at the origin, one is typically interested in the distribution of the process conditioned on the event that absorption has not occurred. In the sequel we

denote by  $\hat{P}_x(t)$  and  $P_x(t)$  the probability density function of the unconditioned and conditioned processes, respectively. Thus for  $x \in \{1, \dots, U\}$  we have

$$P_x(t) = \Pr\{\mathbf{x}(t) = x \mid \mathbf{x}(t) > 0\} = \frac{\hat{P}_x(t)}{1 - \hat{P}_0(t)} \quad (3.9)$$

where  $\hat{P}_x(t) = \Pr\{\mathbf{x}(t) = x\}$ . The limit of  $P_x(t)$  as  $t \rightarrow \infty$  is known as the *quasi-stationary distribution*.

### 3.1.3 Transient distributions

Using the *Kolmogorov equations* for  $\hat{P}_x(t)$  one can show that  $P_x(t)$  satisfies the following differential equations

$$\dot{P}_1 = \chi(2)P_2 - [\eta(1) + \chi(1)]P_1 + P_1\lambda_{\text{ext}}(t) \quad (3.10a)$$

$$\begin{aligned} \dot{P}_x &= \chi(x+1)P_{x+1} - [\eta(x) + \chi(x)]P_x + \eta(x-1)P_{x-1} + P_x\lambda_{\text{ext}}(t), \\ &\quad \forall x \in \{2, 3, \dots, U-1\} \end{aligned} \quad (3.10b)$$

$\vdots$

$$\dot{P}_U = -\chi(U)P_U + \eta(U-1)P_{U-1} + P_U\lambda_{\text{ext}}(t) \quad (3.10c)$$

where  $\lambda_{\text{ext}}(t) := \chi(1)P_1(t)$  [5, 31]. The variable  $\lambda_{\text{ext}}(t)$  is an *extinction rate* in the sense that the conditional probability of extension in an “infinitesimal” interval  $(t, t+dt]$  is given by

$$\Pr\{\mathbf{x}(t+dt) = 0 \mid \mathbf{x}(t) > 0\} = \lambda_{\text{ext}}(t)dt. \quad (3.11)$$

When the population limit  $U$  is small, the above system of equations can be solved numerically. However, for large  $U$ , a more reasonable goal (and one that is of primary interest in applications) is to determine the evolution of some lower-order moments of  $P_x(t)$ .

### 3.2 Time evolution of moments

To model the time evolution of  $\mathbf{x}(t)$ , we consider a Stochastic Hybrid Systems (SHS) with trivial dynamics, two *reset maps*:

$$\mathbf{x} \mapsto \phi_1(\mathbf{x}) := \mathbf{x} + 1, \quad \mathbf{x} \mapsto \phi_2(\mathbf{x}) := \mathbf{x} - 1 \quad (3.12)$$

one corresponding to a birth and the other to a death, with associated *transition intensities* given by

$$\lambda_1(\mathbf{x}) := \eta(\mathbf{x}), \quad \lambda_2(\mathbf{x}) := \chi(\mathbf{x}). \quad (3.13)$$

Given  $m \in \{1, 2, \dots\}$ , we define the  $m^{\text{th}}$  order (*uncentered*) moment for both the unconditioned and conditioned process as

$$\hat{\mu}_m(t) = \sum_{x=1}^{\infty} x^m \hat{P}_x(t) := \mathbf{E}[\mathbf{x}(t)^m], \quad \mu_m(t) = \sum_{x=1}^{\infty} x^m P_x(t), \quad \forall t \geq 0, \quad (3.14)$$

respectively. Using Dynkin's formula for the above SHS, one can conclude that the time evolution of  $\hat{\mu}_m$  is given by

$$\dot{\hat{\mu}}_m = \sum_{h=1}^2 \sum_{r=1}^{m+1} \mathbf{C}_{m+h-r}^m f(m+h-r, h) \hat{\mu}_r, \quad (3.15)$$

where we define  $\mathbf{C}_j^m$  and  $f(j, h)$  as follows<sup>1</sup>  $\forall j, m, h \in \mathbb{N}$ .

$$\mathbf{C}_j^m := \begin{cases} \frac{m!}{(m-j)!j!} & m \geq j \geq 0 \\ 0 & m < j \end{cases} \quad (3.16a)$$

$$f(j, h) := \begin{cases} 0 & j = 0 \\ a_1 + (-1)^j a_2 & j > 0, h = 1 \\ -b_1 + (-1)^j b_2 & j > 0, h = 2. \end{cases} \quad (3.16b)$$

---

<sup>1</sup> $m!$  denotes the factorial of  $m$ .

One can see from the right-hand-side of (3.15), that the time derivative of  $\hat{\mu}_m$  is a *linear combination* of the moments  $\hat{\mu}_r$ , up to order  $r = m + 1$ . Hence the time evolution of the vector  $\hat{\mu} = [\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_k]^T \in \mathbb{R}^k$  is given by

$$\dot{\hat{\mu}} = A\hat{\mu} + B\hat{\mu}_{k+1}, \quad (3.17)$$

for some  $k \times k$  and  $k \times 1$  matrices  $A$  and  $B$  which have the following structure

$$A = \begin{bmatrix} * & * & 0 & 0 & \dots & 0 \\ * & * & * & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ * \end{bmatrix}, \quad (3.18)$$

where each  $*$  denotes a possibly non-zero entry. From (3.9), (3.14) we have

$$\mu_m(t) = \frac{\hat{\mu}_m(t)}{1 - \hat{P}_0(t)}, \quad (3.19)$$

which using (3.17) and  $\dot{\hat{P}}_0(t) = \chi(1)\hat{P}_1(t)$  leads to

$$\dot{\mu} = (A + \lambda_{\text{ext}}(t)I)\mu + B\mu_{k+1}, \quad (3.20)$$

where  $\mu = [\mu_1, \mu_2, \dots, \mu_k]^T \in \mathbb{R}^k$  and  $\lambda_{\text{ext}}(t) := \chi(1)P_1(t)$  is the extinction rate. The dynamics of this system is not closed because the time-derivative of the vector  $\mu$  depends both on the  $k + 1^{\text{th}}$  order moment  $\mu_{k+1}$  and on the extinction rate  $\lambda_{\text{ext}}$ , which are not part of the state  $\mu$ . However, when the mean time to extinction is very large, the perturbation term  $\lambda_{\text{ext}}(t)I$  is very small when compared to the matrix  $A$  and can be ignored [32]. Our goal now is to close the dynamics of (3.20) by approximating  $\mu_{k+1}$  as a nonlinear function of  $\mu$  given by  $\varphi_{k+1}(\mu)$ . This gives the closed approximate moment dynamics

$$\dot{v} = Av + B\varphi_{k+1}(v), \quad v = [v_1, v_2, \dots, v_k]^T. \quad (3.21)$$

We call (3.21) the *truncated moment dynamics* and  $\varphi_{k+1}(\mu)$  the *moment closure function* for  $\mu_{k+1}$ .

### 3.3 Separable derivative-matching moment closure

As done in chapter 1, we seek moment closure functions with the following separable form

$$\varphi_{k+1}^s(\mu) = \prod_{p=1}^k \mu_p^{\gamma_p} \quad (3.22)$$

that match time derivatives between solutions  $\nu$  and  $\mu$ , i.e.,

$$\mu(t_0) = \nu(t_0) \quad \text{and} \quad \frac{d^i \mu(t)}{dt^i} \Big|_{t=t_0} = \frac{d^i \nu(t)}{dt^i} \Big|_{t=t_0}, \quad \forall i \in \{1, \dots\}. \quad (3.23)$$

In the sequel we refer to such  $\varphi_{k+1}^s(\mu)$  as a *Separable Derivative-Matching (SDM) moment closure function* for  $\mu_{k+1}$ . As it is not possible to find  $\gamma_p$  for which (3.23) holds exactly, we relax this condition and simply demand the following

$$\mu(t_0) = \nu(t_0) \quad \text{and} \quad \frac{d^i \mu(t)}{dt^i} \Big|_{t=t_0} = \frac{d^i \nu(t)}{dt^i} \Big|_{t=t_0} + \mathbf{E}[\varepsilon_i(\mathbf{x}(t_0))], \quad (3.24)$$

$\forall i \in \{1, 2, \dots\}$ , where each element of the vector  $\varepsilon_i(\mathbf{x}(t_0))$  is a polynomial in  $\mathbf{x}(t_0)$ .

The following theorem summarizes the main result.

**Theorem 3.1** *Let  $\gamma_p$ ,  $p \in \{1, \dots, k\}$  be chosen as*

$$\gamma_p = (-1)^{k-p} \mathbf{C}_p^{k+1}. \quad (3.25)$$

*Then, for every deterministic initial condition  $\nu(t_0) = \mu(t_0) = [x_0, x_0^2, \dots, x_0^k]^T$ ,  $x_0 \geq 3$  that corresponds to  $\mathbf{x}(t_0) = x_0$  with probability one, we have that*

$$\frac{d\mu(t)}{dt} \Big|_{t=t_0} = \frac{d\nu(t)}{dt} \Big|_{t=t_0} \quad (3.26a)$$

$$\frac{d^2 \mu(t)}{dt^2} \Big|_{t=t_0} = \frac{d^2 \nu(t)}{dt^2} \Big|_{t=t_0} + \varepsilon_2(x_0), \quad (3.26b)$$



where  $\frac{d^i \mu}{dt^i}$  and  $\frac{d^i \nu}{dt^i}$  denote the  $i^{\text{th}}$  time derivative of  $\mu$  and  $\nu$  along the trajectories of the systems (3.20) and (3.21), respectively, and the  $k^{\text{th}}$  element of the vector  $\varepsilon_2(x_0)$  is a polynomial in  $x_0$  of degree 2 with all other elements being zero.

**Proof:** We note from (3.10) that the  $i^{\text{th}}$  time derivative of  $\lambda_{\text{ext}}(t) := \chi(1)P_1(t)$  is a function of  $P_1(t), P_2(t), \dots, P_{i+1}(t)$ . Thus for  $\mathbf{x}(t_0) = x_0$  with probability one, we have that

$$\left. \frac{d^i(\lambda_{\text{ext}}(t)\mu(t))}{dt^i} \right|_{t=t_0} = 0, \quad \forall x_0 \geq i+2. \quad (3.27)$$

Hence choosing  $x_0 \geq 3$  ensures that  $\lambda_{\text{ext}}(t)\mu(t)$  and its time derivative does not appear in the left-hand-side of (3.26).

We have from Theorem 1.2 that  $\gamma_p$  is the solution to the following linear system

$$\mathbf{C}_s^{k+1} = \sum_{p=1}^k \gamma_p \mathbf{C}_s^p, \quad \forall s = \{1, \dots, k\}. \quad (3.28)$$

We now show that solution to (3.28) is unique and given by

$$\gamma_p = (-1)^{k-p} \mathbf{C}_p^{k+1}. \quad (3.29)$$

Towards that end, we observe that for all  $z \in \mathbb{R}$ , one can write using binomial expansion,

$$[1 - (1+z)]^{k+1} = \sum_{p=0}^{k+1} \mathbf{C}_p^{k+1} (-1)^p (1+z)^p = 1 + \sum_{p=1}^{k+1} \mathbf{C}_p^{k+1} (-1)^p \sum_{w=0}^p \mathbf{C}_w^p z^w. \quad (3.30)$$

Equating coefficients for  $z^s$ ,  $s \in \{1, \dots, k\}$  on both sides of (3.30) we have

$$0 = \sum_{p=1}^{k+1} \mathbf{C}_p^{k+1} (-1)^p \mathbf{C}_s^p \Rightarrow (-1)^k \mathbf{C}_s^{k+1} = \sum_{p=1}^k \mathbf{C}_p^{k+1} (-1)^p \mathbf{C}_s^p. \quad (3.31)$$

Comparing (3.31) with (3.28) one can see that a solution to (3.28) will be (3.25). Also the system of  $k$  linear equations (3.28) can be put into the form

$$\mathbf{C} = \Pi \gamma \quad (3.32)$$

where  $\gamma = [\gamma_1, \dots, \gamma_k]^T$ ,  $\mathbf{C} = [\mathbf{C}_1^{k+1}, \dots, \mathbf{C}_k^{k+1}]^T$  and

$$\Pi = \begin{bmatrix} \mathbf{C}_1^1 & \mathbf{C}_1^2 & \dots & \mathbf{C}_1^k \\ 0 & \mathbf{C}_2^2 & \dots & \mathbf{C}_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{C}_k^k \end{bmatrix}. \quad (3.33)$$

As the upper triangular matrix  $\Pi$  is non-singular, the above solution is unique.

Table 1.2 shows the functions  $\varphi_{k+1}^s$  corresponding to  $\gamma_p$  chosen according to (3.25) for  $k = 2, 3$  and 4. The dependence of  $\mu_{k+1}$  on lower order moments  $\mu_1, \dots, \mu_k$  as given by the SDM moment closure function, is consistent with  $\mathbf{x}(t)$  being *lognormally distributed* (see Example 3 in Section 3.4).

### 3.4 Distribution based moment closure

We now review other moment closure techniques for a subsequent comparison with separable derivative-matching moment closure. Most moment closure techniques that appeared in the literature start by assuming a specific class of distributions  $\mathbb{D}$  for the population, and use this assumption to express higher order moments as a function of the lower order ones. We say that such a moment closure function is *consistent with*  $\mathbb{D}$ , which can be formally defined as follows: Let  $\mathbb{D}$  be a class of distributions parameterized by  $w$  parameters  $(q_1, \dots, q_w) \in \mathcal{Q}$ , with the  $m^{th}$  order moment  $\mu_m$  given in terms of the  $q_1, \dots, q_w$  as follows

$$\mu_m = f_m(q_1, \dots, q_w), \quad \forall m \in \{1, 2, \dots\}. \quad (3.34)$$

The moment closure function  $\phi_{k+1}^{\mathbb{D}}(\mu)$  for  $\mu_{k+1}$  is said to be *consistent with the distribution  $\mathbb{D}$*  if, for every  $(q_1, \dots, q_w) \in \mathcal{Q}$ , one has that

$$\mu_{k+1} = f_{k+1}(q_1, \dots, q_w) = \phi_{k+1}^{\mathbb{D}}(\mu) \quad (3.35)$$

where

$$\mu := \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix} = \begin{bmatrix} f_1(q_1, \dots, q_w) \\ \vdots \\ f_k(q_1, \dots, q_w) \end{bmatrix}. \quad (3.36)$$

For well known classes of distributions — such as lognormal, normal, poisson, or binomial — we simply say that  $\phi_{k+1}^{\mathbb{D}}$  is the lognormal, normal, poisson, or binomial moment closure function.

### 3.4.1 Techniques for obtaining $\phi_{k+1}^{\mathbb{D}}$

Generically, when the dimension  $w$  of the parameter space  $\mathcal{Q}$  is the same as the order of truncation  $k$ , the functional equation (3.35)–(3.36) in the “unknown”  $\phi_{k+1}^{\mathbb{D}}(\cdot)$  has a unique solution. In fact, to determine  $\phi_{k+1}^{\mathbb{D}}(\mu)$  one can start by solving (3.36) for  $q_1, \dots, q_w$  in terms of  $\mu_1, \dots, \mu_k$ , and then substituting these back in (3.35) to obtain a unique moment closure function  $\phi_{k+1}^{\mathbb{D}}(\cdot)$ . As we choose  $\mathbb{D}$  to be normal [64], lognormal [22], poisson, or binomial [32], this procedure results in the different moment closure functions shown in Table 3.1.

Difficulties arise when the dimension  $w$  of the parameter space  $\mathcal{Q}$  is strictly smaller than  $k$ , because in this case the functional equation (3.35)–(3.36) does not have a unique solution and one can find infinitely many moment closure functions consistent with the same family of distributions. However, there is a strong incentive to

Table 3.1. Unique Moment Closure Functions for the  $k^{th}$  order truncation ( $k = w$ ) and different distributions  $\mathbb{D}$ .

$\mathbb{D}$	$w$	Unique moment closure function
Normal	2	$\varphi_3^g(\mu) = 3\mu_2\mu_1 - 2\mu_1^3$
Lognormal	2	$\varphi_3^l(\mu) = \frac{\mu_2^3}{\mu_1^3}$
Poisson	1	$\varphi_2^p(\mu) = \mu_1^2 + \mu_1$
Binomial	2	$\varphi_3^b(\mu) = 2\frac{(\mu_2 - \mu_1^2)^2}{\mu_1} - (\mu_2 - \mu_1^2) + 3\mu_1\mu_2 - 2\mu_1^3$

consider this case because, as we shall see shortly, large values for  $k$  generally lead to significantly more accurate moment closures. In the sequel, we illustrate some options for moment closures with  $k > w$ ,

*Example 1:* Consider the class of poisson distributions characterized by their expected value  $\theta$  ( $w = 1$ ). Their moments are given by

$$\mu_1 = f_1(\theta) := \theta \quad (3.37a)$$

$$\mu_2 = f_2(\theta) := \theta(1 + \theta) \quad (3.37b)$$

$$\mu_3 = f_3(\theta) := \theta(1 + 3\theta + \theta^2), \dots \quad (3.37c)$$

[32] proposed the following poisson moment closure function for  $k = 2$ :

$$\varphi_3^{p1}(\mu) = \mu_1 + 3\mu_1\mu_2 - 2\mu_1^3, \quad (3.38)$$

for which it is straightforward to verify that (3.35)–(3.36) holds because

$$\mu_3 = \theta(1 + 3\theta + \theta^2) = \varphi_3^{p1}(\mu), \quad \mu = [\theta, \theta(1 + \theta)]^T. \quad (3.39)$$

However, an alternative choice for the poisson moment closure function that also satisfies (3.35)–(3.36) is given by

$$\phi_3^{p^2}(\mu) = \mu_2 - \mu_1^2 + 3\mu_1\mu_2 - 2\mu_1^3, \quad (3.40)$$

which, as we will see in the next section, performs better than (3.38). The explanation for this lies in the fact that (3.40) has better derivative matching properties than (3.38), in the sense of (3.24). In the sequel we refer to (3.38) and (3.40) as the *Nasell-poisson* and *new-poisson* moment closure function, respectively.

*Example 2:* Consider now the class of normal distributions parameterized by their mean  $\omega$  and variance  $\sigma^2$  ( $w = 2$ ). Their moments are given by

$$\mu_1 = f_1(\theta, \sigma) := \omega \quad (3.41a)$$

$$\mu_2 = f_2(\theta, \sigma) := \omega^2 + \sigma^2 \quad (3.41b)$$

$$\mu_3 = f_3(\theta, \sigma) := \omega(\omega^2 + 3\sigma^2) \quad (3.41c)$$

$$\mu_4 = f_4(\theta, \sigma) := \omega^4 + 6\omega^2\sigma^2 + 3\sigma^4, \dots \quad (3.41d)$$

For  $k = 3$ , any function of the following form is a normal moment closure function for  $\mu_4$

$$\phi_4^g(\mu) := 4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 + h(\mu_1, \mu_2 - \mu_1^2, \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3) \quad (3.42)$$

where  $h(x, y, z)$  is any function with the property that  $h(x, y, 0) = 0$ . To verify that this is so, we note that (3.35)–(3.36) holds because for

$$\mu = [\omega, \quad \omega^2 + \sigma^2, \quad \omega(\omega^2 + 3\sigma^2)]^T, \quad (3.43)$$

we obtain  $h(\mu_1, \mu_2 - \mu_1^2, \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3) = h(\omega, \sigma^2, 0) = 0$  and therefore

$$\phi_4^g(\mu) = \omega^4 + 6\omega^2\sigma^2 + 3\sigma^4 = \mu_4. \quad (3.44)$$

*Example 3:* Finally consider the class of lognormal distributions characterized by the parameters  $\alpha > 0$ ,  $\beta > 0$  ( $w = 2$ ), whose moments are given by

$$\mu_k = f_k(\alpha, \beta) := \alpha^k \beta^{k^2}, \quad \forall k \in \{1, 2, \dots\}. \quad (3.45)$$

For every  $k \geq 2$ , the separable derivative-matching moment closure functions defined by (3.22), with coefficient given by (3.25) in Theorem 3.1 are lognormal moment closure functions for  $\mu_{k+1}$ . We can verify this by noting that (3.35)–(3.36) holds because, from (3.45), (3.25) and (3.28) we conclude that

$$\varphi_{k+1}^s(\mu) = \alpha^{(\sum_{p=1}^k \gamma_p p)} \beta^{(\sum_{p=1}^k \gamma_p p^2)} \quad (3.46a)$$

$$= \alpha^{(\sum_{p=1}^k \gamma_p \mathbf{C}_1^p)} \beta^{(\sum_{p=1}^k \gamma_p \{2\mathbf{C}_2^p + \mathbf{C}_1^p\})}, \quad (3.46b)$$

$$= \alpha^{k+1} \beta^{2\mathbf{C}_2^{k+1} + \mathbf{C}_1^{k+1}} = \alpha^{k+1} \beta^{(k+1)^2} = \mu_{k+1}, \quad (3.46c)$$

where we used the fact that  $2\mathbf{C}_2^{k+1} + \mathbf{C}_1^{k+1} = (k+1)^2$ .

### 3.4.2 Cumulant closure functions

Some literature on moment closure works with moment dynamics expressed in terms of a state vector  $\kappa = [\kappa_1, \dots, \kappa_k]$  where  $\kappa_k(t)$  is the  $k^{th}$  order cumulant<sup>2</sup>, instead of the previously introduced vector  $\mu$  of uncentered moments in (3.20). Then, instead of doing moment closure one performs cumulant closure by approximating  $\kappa_{k+1}$  by a nonlinear function  $\phi_{k+1}(\kappa)$  of  $\kappa_1, \dots, \kappa_k$ , which we refer to as the *cumulant closure function*. The disadvantage of working with  $\kappa$  instead of  $\mu$  is that the dynamics of  $\kappa$

---

<sup>2</sup>The  $k^{th}$  order cumulant,  $\kappa_k$  is given as follows in terms of the uncentered moments

$$\begin{aligned} \kappa_1 &= \mu_1, & \kappa_2 &= \mu_2 - \mu_1^2 \\ \kappa_3 &= \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3, & \kappa_4 &= \mu_4 - 4\mu_1\mu_3 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4, \dots \end{aligned}$$

is always nonlinear. However, for ease of comparison with other papers, we provide in Table 3.2 the cumulant closure functions corresponding to the different moment closure functions discussed so far for  $k = 2$ . We use superscripts  $s$ ,  $l$ ,  $g$ ,  $p1$ ,  $p2$  and  $b$  to denote separable derivative-matching, lognormal, normal, Nasell-poisson, new-poisson and binomial moment closure functions, respectively.

### 3.5 Comparison of transient performance of moment closures

In this section, we compare the transient performance of different moment closure techniques using the error

$$e_{k+1}(t) := \mu_{k+1}(t) - \varphi_{k+1}(\mu(t)) = \sum_{i=0}^{\infty} \frac{(t-t_0)^i}{i!} \varepsilon_{k+1}^i(x_0), \quad (3.48)$$

where

$$\varepsilon_{k+1}^i(x_0) := \frac{d^i \mu_{k+1}(t)}{dt^i} \Big|_{t=t_0} - \frac{d^i \varphi_{k+1}(\mu(t))}{dt^i} \Big|_{t=t_0}. \quad (3.49)$$

We call  $\varepsilon_{k+1}^i(x_0)$  the *derivative matching error*. Ideally, one would like to have  $\varepsilon_{k+1}^i(x_0) = 0$ , but as already pointed out in Section 3.3 this is generally not possible. With deterministic initial conditions as in Theorem 3.1, the derivative matching error is typically a polynomial in  $x_0$ . The lesser the order of this polynomial, the better is  $\varphi_{k+1}(\mu)$  in approximating  $\mu_{k+1}$ .

#### 3.5.1 Moment closures for $k = 2$

We recall from Table 3.2 that  $\varphi_3^l(\mu) = \varphi_3^s(\mu)$ , and therefore we do not need to discuss lognormal moment closure separately. By substituting  $\varphi_3^s(\mu)$ ,  $\varphi_3^g(\mu)$ ,  $\varphi_3^{p1}(\mu)$ ,

Table 3.2. Moment Closure Functions (MCF) for second order truncation ( $k = 2$ ) and corresponding Cumulant Closure Functions (CCF) for  $\mu_3$  and  $\kappa_3$ , respectively, corresponding to the different Moment Closure Techniques (MCT) discussed in this paper. SDM refers to separable derivative-matching.

MCT	MCF	CCF
SDM	$\phi_3^s(\mu) = \frac{\mu_2^3}{\mu_1^3}$	$\phi_3^s(\kappa) = 3\frac{\kappa_2^2}{\kappa_1} + \frac{\kappa_2^3}{\kappa_1^3}$
Normal	$\phi_3^g(\mu) = 3\mu_2\mu_1 - 2\mu_1^3$	$\phi_3^g(\kappa) = 0$
Lognormal	$\phi_3^l(\mu) = \frac{\mu_2^3}{\mu_1^3}$	$\phi_3^l(\mu) = 3\frac{\kappa_2^2}{\kappa_1} + \frac{\kappa_2^3}{\kappa_1^3}$
Nasell-Poisson	$\phi_3^{p1}(\mu) = \mu_1 + 3\mu_1\mu_2 - 2\mu_1^3$	$\phi_3^{p1}(\kappa) = \kappa_1$
New-Poisson	$\phi_3^{p2}(\mu) = \mu_2 - \mu_1^2 + 3\mu_1\mu_2 - 2\mu_1^3$	$\phi_3^{p2}(\kappa) = \kappa_2$
Binomial	$\phi_3^b(\mu) = 2\frac{(\mu_2 - \mu_1^2)^2}{\mu_1} - (\mu_2 - \mu_1^2) + 3\mu_1\mu_2 - 2\mu_1^3$	$\phi_3^b(\mu) = 2\frac{\kappa_2^2}{\kappa_1} - \kappa_2$

$\phi_3^{p2}(\mu)$  and  $\phi_3^b(\mu)$  from Table 3.2 in (3.48)–(3.49), one obtains the corresponding derivative matching errors, which will be denoted using the appropriate superscripts.

Using Table 3.2 and symbolic manipulation in *Mathematica*, we can show that

$${}^s\epsilon_3^0(x_0) = {}^g\epsilon_3^0(x_0) = {}^{p2}\epsilon_3^0(x_0) = {}^b\epsilon_3^0(x_0) = 0 \quad (3.50a)$$

$${}^{p1}\epsilon_3^0(x_0) = -x_0. \quad (3.50b)$$

$${}^*\epsilon_3^i(x_0) \in P_{x_0}(i+1), \quad * = \{s, g, p1, p2, b\}, \quad \forall i \in \{1, 2, \dots\} \quad (3.50c)$$

where  $P_{x_0}(j)$  denotes the set of polynomials in  $x_0$  of degree  $j$ . Since  ${}^{p1}\epsilon_3^0(x_0) = -x_0$ , the Nasell-Poisson moment closure function will have a large initial error, especially



for large initial conditions, when compared to all other moment closure functions. For all  $i \in \{1, 2, \dots\}$ , all of these moment closure functions match derivatives, with the derivative matching error being of the same order in  $x_0$ . The simulation results discussed below show that with the exception of N  sell-Poisson moment closure function, which consistently provides the worst estimates, all other moment closure functions perform fairly well.

*Example:* We consider the stochastic logistic model with

$$a_1 = .30, \quad a_2 = .02, \quad b_1 = .015, \quad b_2 = .001, \quad (3.51)$$

which is used by [27] to model the population dynamics of the African Honey Bee. Using (3.21) with the matrices  $A$  and  $B$  computed in (3.15), we have the following truncated moment dynamics

$$\begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \end{bmatrix} = \begin{bmatrix} 0.28 & -0.016 \\ .32 & .546 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0.032 \end{bmatrix} \varphi_3(\mathbf{v}). \quad (3.52)$$

The time evolution of the moments corresponding to different moment closure techniques is obtained by substituting the appropriate moment closure function from Table 3.2 in place of  $\varphi_3(\mathbf{v})$ . In order to evaluate the performance of these moment closure functions for all time, we compute the exact evolution of the moments  $\mu(t)$ . This is only possible because the population limit  $U = 25$  is small and one can obtain the exact solution by numerically solving the equation (3.10). Figure 3.1 and 3.2 contains plots of the mean and variance errors, respectively, for the different moment closure functions with  $x_0 = 5$  and  $x_0 = 20$ . For  $x_0 = 20$  the binomial moment closure function provides the best estimate both initially and at steady-state, whereas for  $x_0 = 5$  the new-Poisson moment closure function does best initially, but the binomial moment closure function continues to provide the most accurate steady-state estimate. As one

would expect from (3.50), the N  sell-Poisson moment closure function performs the worst.

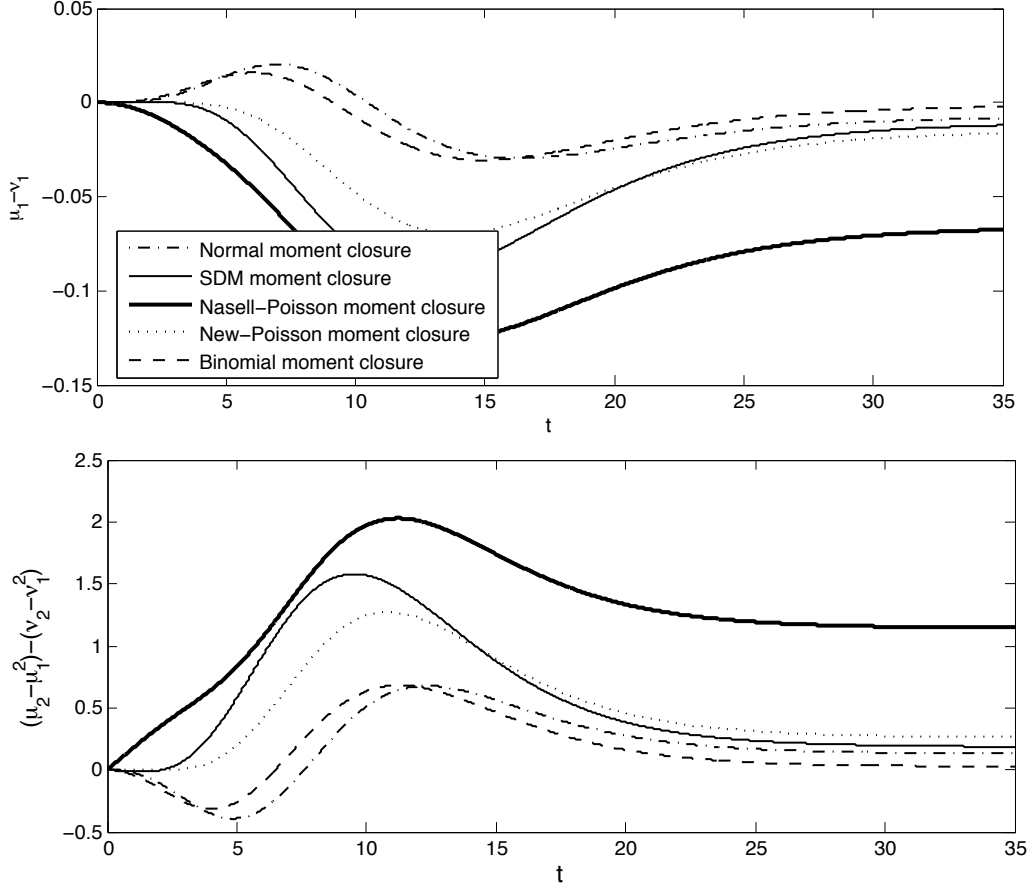


Figure 3.1. Evolutions of the mean error  $\mu_1 - \nu_1$  and of the variance error  $(\mu_2 - \mu_1^2) - (\nu_2 - \nu_1^2)$  for the different moment closure functions in Table 3.2 for  $k = 2$ , with parameters as in (3.51) and  $x_0 = 5$ .

### 3.5.2 Moment closures for $k = 3$

In this section, we propose a new moment closure function given by

$$\varphi_4^z(\mu) = 4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 + \mu_2 - \mu_1^2 \quad (3.53)$$

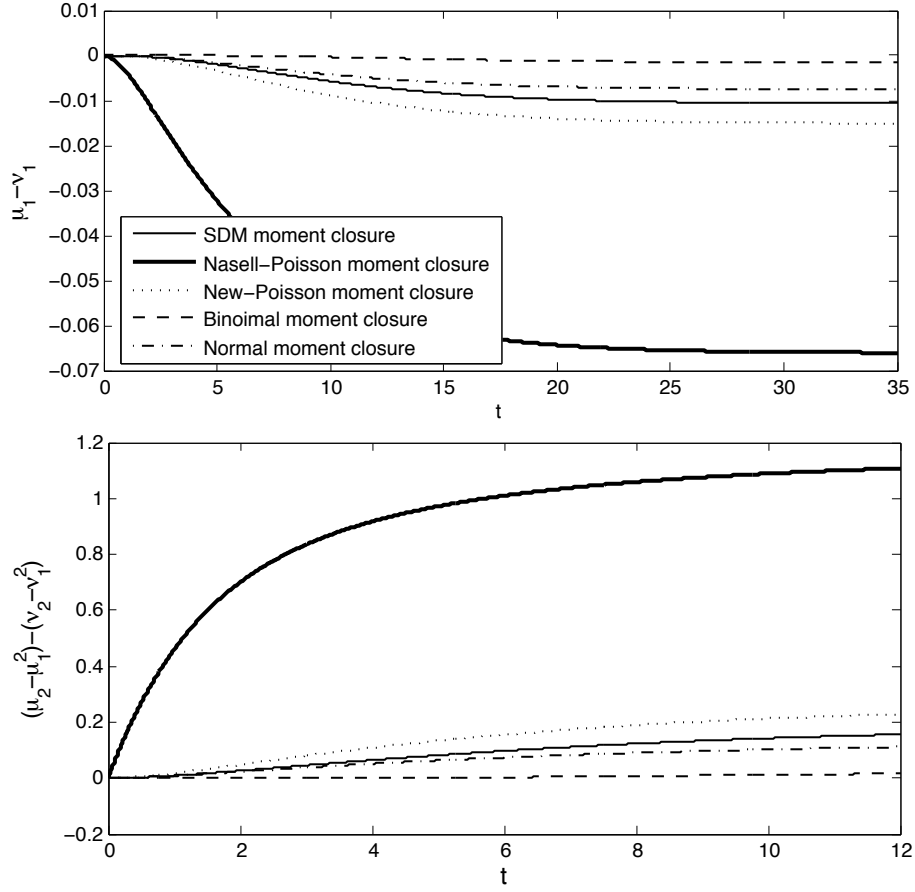


Figure 3.2. Evolutions of the mean error  $\mu_1 - v_1$  and of the variance error  $(\mu_2 - \mu_1^2) - (v_2 - v_1^2)$  for the different moment closure functions in Table 3.2 for  $k = 2$ , with parameters as in (3.51) and  $x_0 = 20$ .

and refer to it as the *Zero first-order error moment closure function*. For comparison purpose we recall from Table 1.2 and Section 3.4 that the SDM and normal moment closure functions for  $k = 3$  are given by

$$\varphi_4^s(\mu) = \frac{\mu_1^4 \mu_3^4}{\mu_2^6} \quad (3.54a)$$

$$\varphi_4^g(\mu) = 4\mu_1 \mu_3 + 3\mu_2^2 - 12\mu_2 \mu_1^2 + 6\mu_1^4, \quad (3.54b)$$

respectively. The above moment closure functions yield the following derivative matching errors:

$${}^*\epsilon_4^0(x_0) = 0, \quad * = \{g, s, z\} \quad (3.55a)$$

$${}^z\epsilon_4^1(x_0) = 0, \quad {}^\dagger\epsilon_4^1(x_0) \in P_{x_0}(2), \dagger = \{g, s\} \quad (3.55b)$$

$${}^*\epsilon_4^i(x_0) \in P_{x_0}(i+1), \quad * = \{g, s, z\}, \quad \forall i \in \{1, 2, \dots\}. \quad (3.55c)$$

From (3.55) one can see the following:

- The normal moment closure function also performs derivative-matching yielding the same order of derivative matching error as the SDM moment closure function, and hence, provides reasonably good estimates for  $\mu_4$ .
- Unlike the other moment closure functions, the zero first-order error moment closure function yields zero  $0^{th}$  and  $1^{st}$  order derivative matching error, and hence, provides the best estimates for  $\mu_4$ , at least near  $t = 0$ .

In order to confirm our predictions above that were based on the expressions (3.55), we consider the stochastic logistic model with parameters as in (3.51),  $k = 3$  and  $x_0 = 20$ . Using (3.15), we have the following truncated moment dynamics

$$\begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \\ \dot{v}_3 \end{bmatrix} = \begin{bmatrix} 0.28 & -0.016 & 0 \\ .32 & .546 & -0.032 \\ .28 & .944 & .798 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0.048 \end{bmatrix} \varphi_3(v). \quad (3.56)$$

Substituting the moment closer functions (3.53)–(3.54b) in place of  $\varphi_3(v)$ , we obtain the corresponding approximate time evolution of moments. Figure 3.3 contains plots of the mean and variance errors. As expected both the normal and SDM moment closure functions provide good estimates. One can also see that the zero first-order

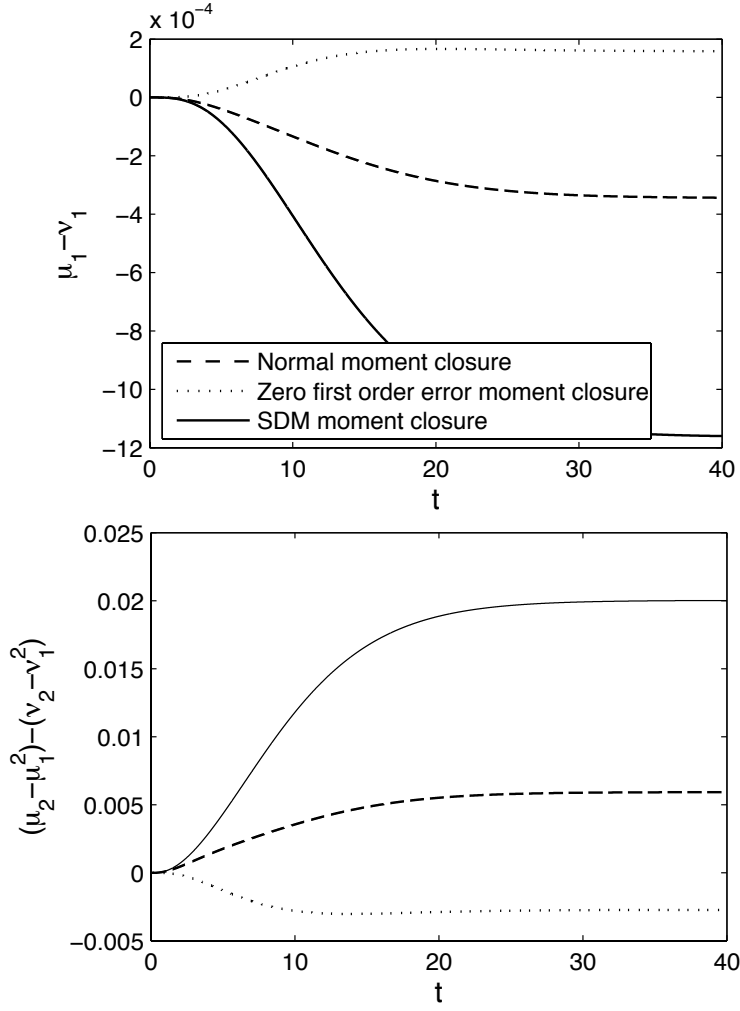


Figure 3.3. Evolutions of the mean error  $\mu_1 - v_1$ , variance error  $(\mu_2 - \mu_1^2) - (v_2 - v_1^2)$  for the different moment closure functions (3.53)–(3.54b) for  $k = 3$ , with parameters as in (3.51) and  $x_0 = 20$ .

error moment closure function (3.53), which guarantees the best approximation near  $t = 0$  actually provides in this case the most accurate estimate for  $\mu_4$  for all time.

Comparing the plots in Figures 3.2 and 3.3, we observe that the mean and variance errors for  $x_0 = 20$  obtained with a third order truncation ( $k = 3$ ) are an order of magnitude smaller than the ones obtained with a second order truncation ( $k = 2$ ).

### 3.6 Steady-state solutions of the truncated moment dynamics

We now show that the SDM moment closure leads to a unique positive equilibrium for the truncated dynamics (3.21) and provide analytical expressions for this equilibrium.

Consider the truncated moment dynamics of degree  $k \geq 2$  with the moment closure functions given in Table 1.2. From (3.21), the steady-state solution  $\mathbf{v}^*$  can be computed by solving

$$0 = A\mathbf{v}^* + B\varphi_{k+1}^s(\mathbf{v}^*) \quad (3.57)$$

where the coefficient of the matrices  $A, B$  can be deduced from (3.15). Solving these equations, we obtain

$$v_2^* = c_1 v_1^* \quad (3.58a)$$

$$\vdots$$

$$v_k^* = c_{n-1} v_1^* \quad (3.58b)$$

$$\varphi_{k+1}^s(\mathbf{v}^*) = c_k v_1^* \quad (3.58c)$$

for appropriate positive real numbers  $c_1, \dots, c_k$ . In terms of the parameters  $a_1, b_1, a_2$  and  $b_2$ , the first three constants  $c_1, c_2, c_3$  are given by the following expressions

$$c_1 = K, \quad c_2 = K^2 + \sigma^2, \quad c_3 = K^3 + 3K\sigma^2 + \bar{\sigma}\sigma^2 \quad (3.59a)$$

$$K = \frac{a_1 - a_2}{b_1 + b_2}, \quad \sigma^2 = \frac{a_1 b_2 + b_1 a_2}{(b_1 + b_2)^2}, \quad \bar{\sigma} = \frac{b_2 - b_1}{b_2 + b_1}. \quad (3.59b)$$

From (3.58) and Table 1.2 we obtain the following steady-state values for  $\varphi_{k+1}^s(\mathbf{v}^*)$ ,

$k \in \{2, 3, 4, \dots\}$ :

$$\varphi_3^s(v^*) = c_1^3, \quad \varphi_4^s(v^*) = \frac{c_2^4}{c_1^6} v_1^{*2}, \quad \varphi_5^s(v^*) = \frac{c_1^{10} c_3^5}{c_2^{10}}. \quad (3.60)$$

Substituting (3.60) in (3.58) yields the following unique non-trivial solutions for the steady-state mean  $v_1^*$

$$v_1^* = \begin{cases} \frac{c_1^3}{c_2} = \frac{K}{1 + \frac{\sigma^2}{K^2}} & k = 2 \\ \frac{c_1^6 c_3}{c_2^4} = \frac{K \left( 1 + \frac{3\sigma^2}{K^2} + \frac{\bar{\sigma}\sigma^2}{K^3} \right)}{\left( 1 + \frac{\sigma^2}{K^2} \right)^4} & k = 3 \\ \frac{c_1^{10} c_3^5}{c_2^{10} c_4} & k = 4. \end{cases} \quad (3.61)$$

The corresponding high-order uncentered moments  $v_2^*, \dots, v_k^*$  can be calculated from (3.58). We conclude that the SDM moment closure function always yields a unique non-trivial positive, real, steady-state for every truncation order  $k \geq 2$ . Moreover, the separable structure of the SDM moment closure leads to analytical expressions for the steady-state moments. In contrast, finding the steady-state moments for the normal moment closure requires solving an  $k^{th}$  degree polynomial in  $v_1^*$  and then identifying the biologically relevant steady-state among the  $k$  roots of the polynomial. For  $k > 2$  this can generally only be done numerically and one does not obtain analytic expressions for the steady-state moments [25].

## 3.7 Conclusion

A procedure for constructing moment closures for the stochastic logistic model was presented. This was done by first assuming a separable form for the moment closure function  $\varphi_{k+1}(v)$ , and then, matching its time derivatives with  $\mu_{k+1}$ , at some initial time  $t_0$  for a basis of initial conditions  $x(t_0) = x_0$ . We showed that there exists a

unique separable derivative-matching moment closure function for which the  $i^{th}$  order derivative matching error is a polynomial in  $x_0$  of degree  $i + 1$  for all  $i \in \{1, 2, \dots\}$  and zero for  $i = 0$ . Explicit formulas to construct these moment closure functions for arbitrary order of truncation  $k$  were provided with higher values of  $k$  leading to better approximation of the actual moment dynamics.

The separable structure of this moment closure greatly simplified the process of finding the steady-state of the truncated moment dynamics which were always unique, real and positive. Comparisons with alternative moment closure techniques available in literature illustrated how derivative matching can be used as a effective tool for gauging the performance of moment closure functions. We showed that for  $k = 2$ , with the exception of the N  sell-Poisson, all other moment closure functions in Table 3.2 perform derivative matching and provide fairly good estimates for  $\mu_3$ . For  $k = 3$ , a new zero first-order error moment closure function was also proposed, guaranteeing better approximations, at least locally in time, as compared to the other moment closure techniques discussed in this chapter.

### 3.7.1 Future work

The truncated moment dynamics presented in this chapter only capture the quasi-stationary distribution and do not provide information about the time taken to reach extinction. Finding alternative moment closure techniques that provide information about extinction is a subject for future research.



## **Chapter 4**

# **Optimal feedback strength for noise suppression in auto-regulatory gene networks**

Gene expression and regulation is inherently a noisy process. The origins of this stochasticity lie in the probabilistic nature of transcription and translation and low copy numbers of RNAs and proteins within cells, which can lead to large statistical fluctuations in molecule numbers. Recent work [61, 4, 8, 12, 44] has provided considerable experimental evidence for these stochastic fluctuations and may explain for the large amounts of cell to cell variation observed in genetically identical cells exposed to the same environmental conditions [53, 28]. Various gene network motifs within cells decrease/increase these stochastic fluctuations. A common such motif is an auto-regulatory gene network where the protein expressed from the gene inhibits/activates its own transcription [3, 60]. Both theoretical and experimental studies have shown

that negative feedback in these auto-regulatory gene networks reduces stochastic fluctuations in the protein population [46, 35, 57, 6, 59, 48] whereas positive feedback has the opposite effect [17, 9].

Auto-regulatory gene networks are characterized by their *transcriptional response*  $g(\mathbf{x})$ , which determines the transcription rate of the gene as a non-linear function  $g$  of the protein molecular count  $\mathbf{x}$  within the cell. Monotonic decreasing and increasing functions  $g(\mathbf{x})$  denote negative and positive feedback, respectively. The noise in the protein population is quantified by its *coefficient of variation* defined as the ratio of the standard deviation to the average number of protein molecules. Previous work has shown that this protein noise level is determined by a combination of two components [56, 37]. The first is the intrinsic noise, which represents the stochastic fluctuations in the protein population arising due to random protein formation and degradation events. The second component is the extrinsic noise, which corresponds to fluctuations in the protein numbers arising due to an exogenous noise source driving the auto-regulatory gene network, for example, fluctuations in gene copy numbers, enzyme levels, and/or environmental stimuli. Table 4.1 provides a summary of the notations used for the different forms of noise in the protein population.

Our goal is to understand how these different components of protein noise can be modulated by manipulating the *response time* of the auto-regulatory gene network, which is defined as follows: assuming  $\mathbf{x}^*$  to be the steady-state average protein count, the *response time*  $T_r$  is the time taken for any initial perturbation about  $\mathbf{x}^*$  to decay by 50% of its initial value. Negative and positive feedback in the auto-regulatory gene network, decreases and increases the response time, respectively, from its value when there is no feedback (i.e., when the transcriptional response  $g(\mathbf{x})$  is a constant and independent of  $\mathbf{x}$ ).

Table 4.1. A summary of the notation used in this chapter. All estimates of noise are based on a linear approximation of the transcriptional response.

$CV_{tot}$	Total noise in protein numbers
$CV_{ext}$	Extrinsic noise in protein numbers
$CV_{int}$	Intrinsic noise in protein numbers
$CV_z$	Noise in the exogenous signal driving the gene network.
$CV_{tot-nr}$	Total noise in protein numbers when there is no feedback
$CV_{ext-nr}$	Extrinsic noise in protein numbers when there is no feedback
$CV_{int-nr}$	Intrinsic noise in protein numbers when there is no feedback
$CV_{tot-min}$	Minimum possible total noise in protein numbers with optimal negative feedback
$CV_{ext-min}$	Minimum possible extrinsic noise in protein numbers with optimal negative feedback
$CV_{int-min}$	Minimum possible intrinsic noise in protein numbers with optimal negative feedback

We consider a simple model of gene expression where each expression event produces a random number of protein molecules according to a given probability distribution. Details on the stochastic formulation of this model are provided in Section 4.1. In Section 4.2, we determine the intrinsic noise in the protein population. Using a linear approximation for the transcriptional response

$$g(\mathbf{x}) \approx g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*), \quad (4.1)$$

where  $\mathbf{x}^*$  is the steady-state average protein count, we show that the intrinsic noise level is proportional to the ratio  $T_r/\mathbf{x}^*$ . Hence for a fixed  $\mathbf{x}^*$ , decreasing the protein's response time  $T_r$  attenuates the intrinsic noise whereas increasing the response time magnifies it. We also investigate the effects of non-linearities in the transcriptional response and show that a concave (convex) transcriptional response causes the noise in the protein to be smaller (larger) than what would be predicted by the linear transcriptional response in (4.1).

We next quantify the extrinsic noise in the protein population. In Section 4.3, we derive analytical formulas that decompose the total noise in the protein into its extrinsic and intrinsic components. These formulas are simple generalizations of previous work, notably that of [37], where the number of protein molecules produced per expression event was not random but deterministic and equal to one. We show that for a given decrease in  $T_r$  through negative feedback, the extrinsic noise decreases by a much larger amount than does the intrinsic noise. Thus negative feedback is much more effective in reducing the extrinsic component of protein noise than its intrinsic component.

In Section 4.4, we use the above results to quantify noise in auto-regulatory gene networks that involve a common negative feedback with transcriptional response given

by

$$g(\mathbf{x}) = g_0 \left( b + \frac{1-b}{1+(a\mathbf{x})^M} \right), \quad 0 \leq b < 1 \quad (4.2)$$

where  $M \geq 1$  denotes the *Hill coefficient* and  $g_0$  corresponds to the transcription rate when there is no feedback (i.e.,  $a = 0$ ) [3, 58]. The positive constant  $b$  is less than one and is chosen such that the product  $g_0 b$  represents the minimum level of transcription rate obtained in the limit of a large population  $\mathbf{x} \rightarrow \infty$ . The constant  $a$  characterizes the *feedback strength* and is determined by the binding affinity of the protein to the promoter of the gene.

We perform a systematic analysis of how the protein noise level changes as the feedback strength  $a$  is increased from an initial value of zero. We first consider the situation where extrinsic noise is absent or negligible and intrinsic noise dominates the total noise in the protein population. In such a scenario, we show that if the Hill coefficient is close to one, then the protein noise level actually increases as we increase the feedback strength. However, for Hill coefficients larger than one, the protein noise level first decreases as we increase the feedback strength from zero and achieves a minimum value at some optimal level of feedback strength. Increasing the feedback strength above this optimal value causes an increase in the noise level. In summary, for Hill coefficients  $M$  larger than one, we obtain a U-shaped noise profile as the feedback strength is increased. We quantify both the optimal level of feedback strength and the *limit of noise suppression*, which is defined as the ratio of the minimum possible noise in the protein population to the protein noise level when there is no feedback (i.e.,  $a = 0$ ). When the intrinsic noise dominates the total noise in the protein population, this limit is given by the simple expression

$$\sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} \leq 1 \quad (4.3)$$

which essentially only depends on the Hill coefficient, as typically  $b$  is much smaller than one.

We next consider the situation where the extrinsic noise is not negligible and both extrinsic and intrinsic noise are present. We show that in this case, irrespective of the value of the Hill coefficient, the protein noise level always follows a U-shaped profile as the feedback strength is increased. This means that the noise level is minimized at some optimal value of feedback strength and decreasing or increasing feedback strength away from this optima will always causes an increase in the noise level. We again provide analytical formulae for the limit of noise suppression and show that in this case (when extrinsic noise is not small compared to the intrinsic noise), this ratio is much lower than what is given by (4.3), which corresponds to the situation where there is no extrinsic noise. In fact, we show that determining how much the limit of noise suppression deviates from (4.3) can be used to estimate how much extrinsic noise is present in the gene network.

In Section 4.5 we validate our theoretical results by using experimental data from a synthetic auto-regulatory gene network described in [11]. As predicted, we indeed see a U-shaped profile for the protein noise level as the feedback strength is experimentally manipulated. We also explain observations in [11] showing that for small levels of extrinsic noise, no U-shaped profile is observed, and instead, the protein noise level monotonically increase as the feedback strength is increased. Finally, we illustrate how the experimentally determined limit of noise suppression can be used to estimate the noise in the exogenous signal. Matching these estimates with independent measurements of noise associated with the plasmid population, we confirmed that variability in plasmid numbers was the major source of extrinsic noise in the synthetic gene network in [11].

## 4.1 Un-regulated gene expression

We consider a simple model of gene expression where a gene expresses a protein  $X$  in bursts that occur at a rate  $K_x$ . Each expression event leads to the formation of  $\mathbf{N}_x$  molecules of the protein  $X$ . Recent work suggests that the burst of proteins from each mRNA transcript follows a geometric distribution [24]. Thus instead of assuming  $\mathbf{N}_x$  to be a constant, we assume it to be a random variable with mean  $N_x$  and variance  $V_x^2$ . We also assume that the protein decays at a constant rate  $d_x$ . Our model omits the mRNA dynamics. This is a valid approximation as long as the protein's life time is much longer than the mRNA's life time, which is generally the case in gene-protein networks [38] (see Appendices (B.5) and (B.6) for stochastic models of gene expression that consider mRNA dynamics). Ignoring the mRNA dynamics leads to relatively simple expressions for the protein noise level, which help develop a qualitative understanding of how noise level changes in response to alterations of the gene network parameters.

In a stochastic formulation, gene expression and protein degradation are treated as probabilistic events with probabilities of occurring in an infinitesimal time interval  $(t, t + dt]$  given by

$$\Pr\{\mathbf{x}(t + dt) = x + \mathbf{N}_x \mid \mathbf{x}(t) = x\} = K_x dt \quad (4.4a)$$

$$\Pr\{\mathbf{x}(t + dt) = x - 1 \mid \mathbf{x}(t) = x\} = d_x x dt, \quad (4.4b)$$

respectively, where  $\mathbf{x}(t)$  denotes the number of molecules of protein  $X$  at time  $t$ .

A convenient way to model the time evolution of the number of molecules  $\mathbf{x}$  is through a Stochastic Hybrid System (SHS) characterized by trivial continuous dynamics

$$\dot{\mathbf{x}} = 0, \quad (4.5)$$

and two reset maps

$$\mathbf{x} \mapsto \phi_1(\mathbf{x}) = \mathbf{x} + \mathbf{N}_x, \quad \mathbf{x} \mapsto \phi_2(\mathbf{x}) = \mathbf{x} - 1 \quad (4.6)$$

with corresponding transition intensities given by

$$\lambda_1(\mathbf{x}) = K_x, \quad \lambda_2(\mathbf{x}) = d_x \mathbf{x} \quad (4.7)$$

[19]. In order to gauge the noise level in the protein population, we determine the time evolution of the first and second order moments of  $\mathbf{x}$ , i.e., the expected values  $\mathbf{E}[\mathbf{x}]$  and  $\mathbf{E}[\mathbf{x}^2]$ . The moment dynamics can be obtained using the Dynkin's formula for the above SHS, according to which, for every differentiable function  $\psi(\mathbf{x})$  we have that

$$\frac{d\mathbf{E}[\psi(\mathbf{x})]}{dt} = \mathbf{E} \left[ \sum_{i=1}^2 (\psi(\phi_i(\mathbf{x})) - \psi(\mathbf{x})) \lambda_i(\mathbf{x}) \right] \quad (4.8)$$

[10, 18]. Taking  $\psi(\mathbf{x}) = \mathbf{x}$  and  $\psi(\mathbf{x}) = \mathbf{x}^2$  in (4.8) we obtain the following moment dynamics

$$\frac{d\mathbf{E}[\mathbf{x}]}{dt} = N_x K_x - d_x \mathbf{E}[\mathbf{x}], \quad (4.9a)$$

$$\frac{d\mathbf{E}[\mathbf{x}^2]}{dt} = K_x(N_x^2 + V_x^2) + d_x \mathbf{E}[\mathbf{x}] + 2K_x N_x \mathbf{E}[\mathbf{x}] - 2d_x \mathbf{E}[\mathbf{x}^2]. \quad (4.9b)$$

As  $t \rightarrow \infty$ , the first and second order moments converge to constant steady-state values given by

$$\mathbf{x}^* := \lim_{t \rightarrow \infty} \mathbf{E}[\mathbf{x}(t)] = \frac{N_x K_x}{d_x} \quad (4.10a)$$

$$\mathbf{E}^*[\mathbf{x}^2] := \lim_{t \rightarrow \infty} \mathbf{E}[\mathbf{x}^2(t)] = \frac{K_x d_x N_x + 2K_x^2 N_x^2 + K_x d_x (N_x^2 + V_x^2)}{2d_x^2}. \quad (4.10b)$$

We quantify the noise in  $\mathbf{x}(t)$  by its *coefficient of variation* defined as the ratio of the standard deviation in protein numbers to the average number of protein molecules.

Using the above steady-state values we obtain

$$CV_{int-nr}^2 = \frac{\mathbf{E}^*[\mathbf{x}^2] - \mathbf{x}^{*2}}{\mathbf{x}^{*2}} = \frac{d_x(N_x^2 + V_x^2 + N_x)}{2K_x N_x^2} = \frac{(N_x^2 + V_x^2 + N_x)}{2\mathbf{x}^* N_x}. \quad (4.11)$$



This formula quantifies the noise in the protein  $X$  solely due to random gene expression and protein degradation, and is referred to as the *intrinsic noise* in the protein population when there is no regulation. Note that the noise in the protein increases with the variance  $V_x^2$  in the number of protein molecules produced in each transcription event.

A well known special case of (4.11) is obtained for  $N_x = 1$  and  $V_x = 0$ , for which  $\mathbf{x}(t)$  has a Poisson distribution and  $CV_{int-nr}^2 = 1/\mathbf{x}^*$ . In the next section we examine what happens to this intrinsic noise when the gene expression rate is not a constant but a function of the number of protein molecules.

## 4.2 Auto-regulatory gene expression

Often the expressed protein binds to the promoter region of its own gene. In doing so it either recruits the enzyme RNA Polymerase to the promoter (which leads to an increase in gene expression) or blocks RNA Polymerase from binding to the promoter (which causes a decrease in gene expression). Such gene expression is referred to as an *auto-regulatory gene network*. We model this network by assuming that the rate of gene expression is no longer a constant and is instead a function  $g(\mathbf{x})$  of the number of protein molecules  $\mathbf{x}$  (as shown in Figure 4.1). We refer to the function  $g(\mathbf{x})$  as the *transcriptional response* of the network. This transcriptional response can be formally derived assuming that the rate of binding and dissociation between the protein and its promoter is much faster than the dynamics of protein production and degradation [3] or it can be determined directly from experiments. Monotonic decreasing and increasing functions  $g(\mathbf{x})$  denote negative and positive feedback, respectively.

When an auto-regulation mechanism is present, the probabilities of gene expres-

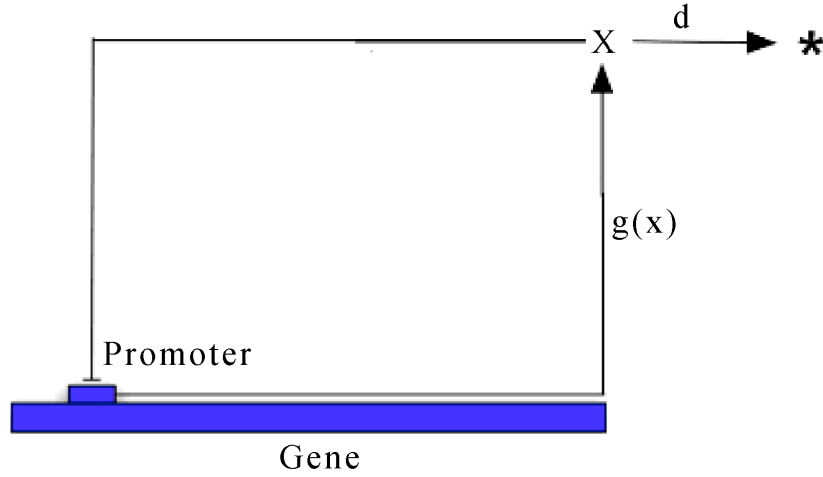


Figure 4.1. An auto-regulatory gene network.

sion and protein degradation events occurring in an infinitesimal time interval  $(t, t + dt]$  are given by

$$\Pr\{\mathbf{x}(t + dt) = x + \mathbf{N}_x \mid \mathbf{x}(t) = x\} = g(x)dt \quad (4.12a)$$

$$\Pr\{\mathbf{x}(t + dt) = x - 1 \mid \mathbf{x}(t) = x\} = d_x x dt. \quad (4.12b)$$

To write the moment dynamics of  $\mathbf{x}$  we first approximate  $g(\mathbf{x})$  by a polynomial in  $\mathbf{x}$ , which is done by expanding  $g(\mathbf{x})$  as a Taylor series expansion

$$g(\mathbf{x}) = g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}g''(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)^2 + \dots, \quad (4.13)$$

about the steady-state average number of protein molecules  $\mathbf{x}^*$ .

### 4.2.1 Linear transcriptional response

For now, we ignore quadratic and higher order terms in (4.13), which results in a linear transcriptional response

$$g(\mathbf{x}) \approx g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*). \quad (4.14)$$

This approximation is valid as long as the stochastic fluctuations in the protein do not leave the region in which  $g(\mathbf{x})$  is approximately linear. As in Section 4.1, we model the time evolution of  $\mathbf{x}$  through a Stochastic Hybrid System (SHS) but now the transition intensities are given by  $\lambda_1(\mathbf{x}) = g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$  and  $\lambda_2(\mathbf{x}) = d_x \mathbf{x}$ . Using the Dynkin's formula for this modified SHS we obtain the following dynamics for the mean  $\mathbf{E}[\mathbf{x}]$ :

$$\frac{d\mathbf{E}[\mathbf{x}]}{dt} = N_x g(\mathbf{x}^*) - d_x \mathbf{x}^* + (N_x g'(\mathbf{x}^*) - d_x)(\mathbf{E}[\mathbf{x}] - \mathbf{x}^*) \quad (4.15)$$

and therefore the steady-state value  $\mathbf{x}^*$  for the mean population  $\mathbf{E}[\mathbf{x}]$  must satisfy

$$N_x g(\mathbf{x}^*) = d_x \mathbf{x}^*. \quad (4.16)$$

To be biologically meaningful, the average  $\mathbf{E}[\mathbf{x}]$  must remain bounded which means that the linear system given by (4.15) must have a negative eigenvalue

$$\lambda := N_x g'(\mathbf{x}^*) - d_x < 0. \quad (4.17)$$

This eigenvalue  $\lambda$  can be expressed in terms of the response time  $T_r$  of the protein, a quantity that can be measured experimentally. The response time  $T_r$  is defined as the time taken for  $\mathbf{E}[\mathbf{x}(t)] - \mathbf{x}^*$  to decay by 50% of its initial condition, i.e.,  $\mathbf{E}[\mathbf{x}(T_r)] - \mathbf{x}^* = (\mathbf{E}[\mathbf{x}(0)] - \mathbf{x}^*)/2$  and is given by

$$T_r = -\frac{\ln(2)}{\lambda} > 0, \quad \lambda := N_x g'(\mathbf{x}^*) - d_x < 0. \quad (4.18)$$

Negative feedback, which correspond to  $g'(\mathbf{x}^*) < 0$ , decreases the response time from the value  $T_{nr} = \ln(2)/d_x$  that corresponds to the absence of feedback (i.e.,  $g'(\mathbf{x}^*) = 0$ ). Positive feedback has an opposite effect.

We now compute the coefficient of variation of  $\mathbf{x}(t)$  by writing the moment dynamics for the second order moment  $\mathbf{E}[\mathbf{x}^2]$ . Using (4.8), with  $\psi(\mathbf{x}) = \mathbf{x}^2$  we obtain the

following time derivative for  $\mathbf{E}[\mathbf{x}^2]$ :

$$\begin{aligned} \frac{d\mathbf{E}[\mathbf{x}^2]}{dt} = & [g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)](N_x^2 + V_x^2) + d_x \mathbf{E}[\mathbf{x}] + 2[g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)]N_x \mathbf{E}[\mathbf{x}] \\ & - 2d_x \mathbf{E}[\mathbf{x}^2] + g'(\mathbf{x}^*)(N_x^2 + V_x^2)\mathbf{E}[\mathbf{x}] + 2g'(\mathbf{x}^*)N_x \mathbf{E}[\mathbf{x}^2]. \end{aligned} \quad (4.19)$$

Performing a steady-state analysis of the above equations and using (4.16) we obtain the following steady-state coefficient of variation

$$CV_{int} = \sqrt{\frac{d_x(N_x^2 + V_x^2 + N_x)}{2IN_x^2}}, \quad I = g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*) \quad (4.20)$$

where  $I$  can be interpreted as the y-intercept of the tangent to the transcriptional response  $g(\mathbf{x})$  at  $\mathbf{x} = \mathbf{x}^*$  (see Figure 4.2). Using (4.16), (4.18), and (4.20) we can also relate the intrinsic noise to the response time  $T_r$  of the protein as

$$CV_{int} = \sqrt{\frac{T_r}{T_{nr}} \frac{N_x^2 + V_x^2 + N_x}{2\mathbf{x}^* N_x}} \quad (4.21)$$

where  $T_{nr} = \ln(2)/d_x$  is the protein's response time when there is no regulation in gene expression (i.e.,  $g'(\mathbf{x}^*) = 0$  and the transcription rate is a constant as in Section 4.1).

The formula in (4.21) shows that the intrinsic noise level in auto-regulatory gene networks is determined by three factors: the average number of protein molecules  $\mathbf{x}^*$ , the response time of the protein  $T_r$  and the gene expression burst characteristics, i.e.,  $N_x$  and  $V_x^2$ . From (4.20) we also conclude that for a fixed  $\mathbf{x}^*$ , making the slope  $g'(\mathbf{x}^*)$  more negative causes a decrease in the response time and leads to attenuation of intrinsic noise in the protein population. However, as we will see later, experimental manipulations that change the response time typically also alter  $\mathbf{x}^*$ , in which case, attenuation or magnification of intrinsic noise will depend on whether the ratio  $T_r/\mathbf{x}^*$  in (4.21) decreases or increases, respectively.

When the number of proteins produced per mRNA follows a geometric distribution

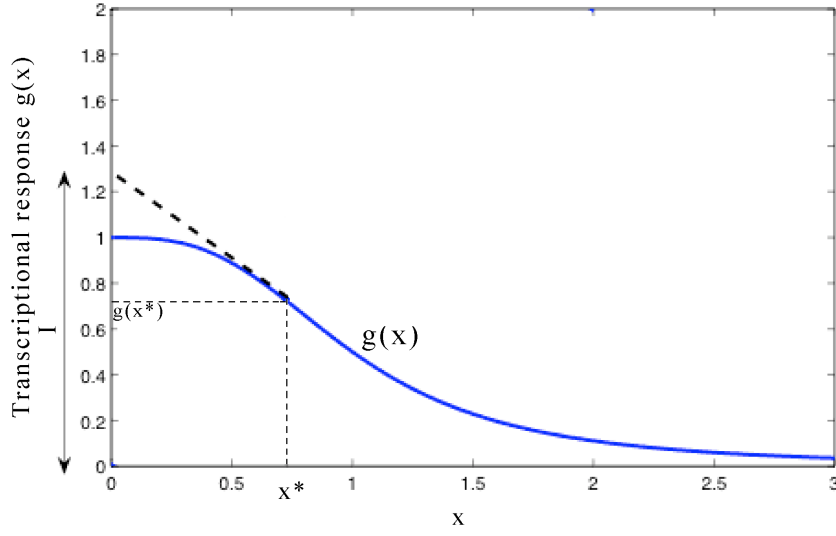


Figure 4.2. A graphical interpretation of the quantity  $I = g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)$  in (4.20) for any arbitrary transcriptional response  $g(\mathbf{x})$ :  $I$  is the intercept of the tangent to the transcriptional response  $g(\mathbf{x})$  at  $\mathbf{x} = \mathbf{x}^*$  with the y-axis.

[24], the variance  $V_x^2$  is equal to  $N_x^2 - N_x$ . In this case (4.21) simplifies to

$$CV_{int} = \sqrt{\frac{T_r N_x}{T_{nr} \mathbf{x}^*}} \quad (4.22)$$

which shows that for all other parameters fixed, the intrinsic noise increases as we increase the average number  $N_x$  of proteins produced per gene expression event, which is consistent with other theoretical and experimental observations [36, 58].

An important feature of (4.22) is that it relates the noise in the protein to parameters that can be experimentally determined. In particular, response times can be measured by tracking the time evolution of the number of molecules within the cell and  $N_x = L_x/a_x$  where  $L_x$  is the translation rate of the mRNA and  $a_x$  is the mRNA degradation rate. For example, in [45] an auto-regulatory gene network was designed where the protein repressed its own transcription. The protein was fluorescently tagged which

allowed one to compute the time evolution of the average number of protein molecules in the cell. Figure 4.3 plots this time evolution with and without negative feedback

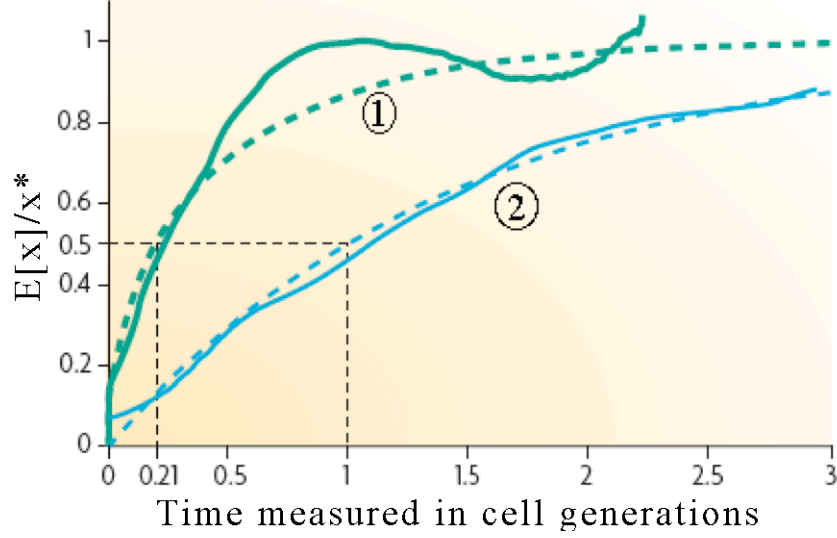


Figure 4.3. Time evolution of the average number of protein molecules. 1) corresponds to the case when there is negative feedback (i.e., protein repressed its own transcription) and 2) corresponds to the case when there is gene expression with no negative feedback. The solid and dashed lines represent experimentally measured and fitted approximations to the time evolution of the average number of protein molecules, respectively. This figure was taken from [3].

in the gene. The promoter strength was appropriately adjusted such that the steady-state population  $\mathbf{x}^*$  of the protein was the same in both cases. The figure shows that with negative feedback it takes about  $T_r = 0.21$  time unit for the protein count to reach half of its steady-state average protein count  $\mathbf{x}^*$ . The response time when there is no feedback is  $T_{nr} = 1$  time unit, which is five times larger than  $T_r$ . We can then conclude from (4.22) that for this network, the presence of negative feedback reduces the intrinsic noise levels in the protein population by a factor of  $\sqrt{1/0.21} \approx 2.2$ .

### 4.2.2 Lambda repressor gene network

We now use the results of the previous sections to investigate an important gene motif that arises in a gene associated with lambda phage, a virus that infects bacteria. The lambda phage has a gene that encodes for a protein called the lambda repressor, which activates its own transcription, resulting in positive feedback. Large levels of this protein causes the virus to lysogenize (i.e., integrate its own chromosome into the bacteria DNA). The transcriptional response of an auto-regulatory gene network with positive feedback is typically given by

$$g_1(\mathbf{x}) = g_0 \left( b + \frac{1-b}{1 + (\alpha\mathbf{x})^M} \right), \quad 1 < b \quad (4.23)$$

where  $\alpha, \beta$  are positive constants and  $M \geq 1$  represents the Hill coefficient [2]. This function is a sigmoidal-shaped monotonically increasing function, such as the transcriptional response  $g_1(\mathbf{x})$  in Figure 4.4. However, for the lambda repressor gene the transcriptional response has been modified and the protein activates the gene only when the number of protein molecules is small. At larger protein populations the protein inhibits its own transcription [42]. As a consequence, the transcriptional response of this particular gene network is an increasing function when  $\mathbf{x}$  is small and a decreasing function when  $\mathbf{x}$  becomes large, as the modified transcriptional response  $g_2(\mathbf{x})$  in Figure 4.4.

As can be seen in Figure 4.4, the modified transcriptional response  $g_2(\mathbf{x})$  has a larger intercept  $I_2$  than the original transcriptional response  $g_1(\mathbf{x})$ . Consequently, in view of (4.20) the modified transcriptional response leads to smaller levels of intrinsic noise in the protein than the original transcriptional response. Low stochastic fluctuations in the lambda repressor population ensure that its number do not become small just by random chance, which will cause the virus to come out of lysogeny and lyse

the cell.

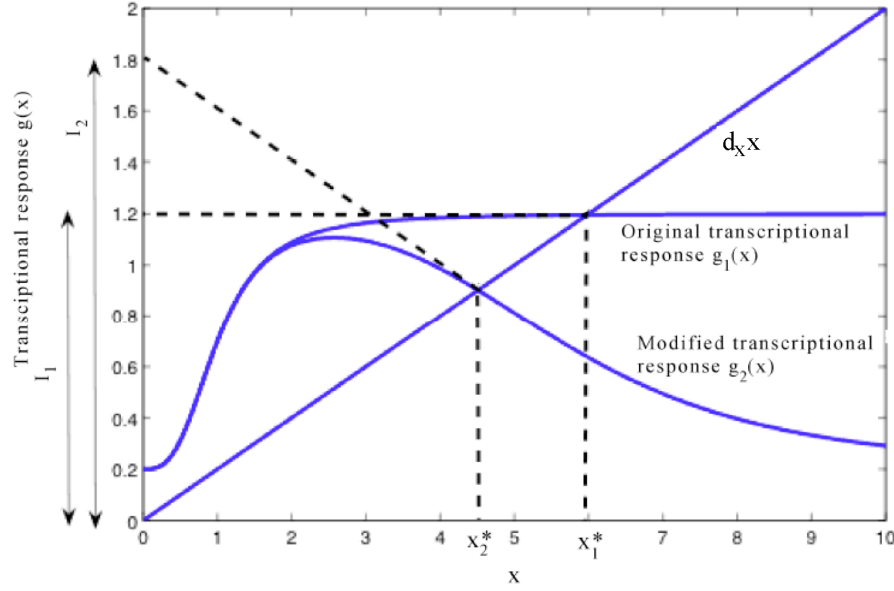


Figure 4.4.  $g_1(\mathbf{x})$  is the standard transcriptional response of a gene network with positive feedback while  $g_2(\mathbf{x})$  is the observed transcriptional response in case of the gene in lambda phage encoding the protein lambda repressor.  $I_1$  and  $I_2$  is the y-intercept of the tangent to the corresponding transcriptional response  $g(\mathbf{x})$  at  $(\mathbf{x}^*, g(\mathbf{x}^*))$ .

### 4.2.3 Effect of nonlinearities

We now examine the effects of non-linear quadratic terms in  $g(\mathbf{x})$ . Towards this end we approximate  $g(\mathbf{x})$  as the following second order polynomial

$$g(\mathbf{x}) = g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}g''(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)^2 \quad (4.24)$$

where  $\mathbf{x}^*$  is given by (4.16), still ignoring cubic and higher order terms in  $\mathbf{x} - \mathbf{x}^*$ . Referring the reader to Appendix B.1 for more details, the intrinsic noise in the protein



population can now be approximated by

$$CV_{int-quad}^2 = \frac{CV_{int}^2}{1 + \frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV_{int}^2}{2\lambda}}, \quad \lambda := N_x g'(\mathbf{x}^*) - d_x < 0 \quad (4.25)$$

where  $CV_{int}$  is the intrinsic noise when  $g(\mathbf{x})$  was assumed linear and is given by (4.20). To obtain (4.25) we assumed that  $CV_{int-quad}^2$  is much smaller than one and the distribution of the protein population is symmetrically distributed about its mean. The above result shows two important points. First, a transcriptional response that is concave at  $\mathbf{x}^*$  ( $g''(\mathbf{x}^*) < 0$ ) results in smaller intrinsic noise than what is predicted by  $CV_{int}$ , whereas a convex response ( $g''(\mathbf{x}^*) > 0$ ) has the opposite effect. Second, as long as  $CV_{int}$  and the non-linearity in the transcriptional response are small in the sense that

$$\left| \frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV_{int}^2}{2\lambda} \right| \ll 1, \quad (4.26)$$

linearizing the transcriptional response will yield a good approximation for the intrinsic noise in the protein population.

### 4.3 Extrinsic and intrinsic contributions to noise

We now consider extrinsic noise in the protein population arising due to an exogenous noise source driving the auto-regulatory gene network. Towards that end, we consider a transcriptional response  $g(\mathbf{x}, \mathbf{z})$  that also depends on a noisy exogenous signal  $\mathbf{z}$ .

The transcriptional response  $g(\mathbf{x}, \mathbf{z})$  may take different forms. For example, if the gene is encoded on a low-copy plasmid, then fluctuations in the number of copies of the plasmid are known to be a major source of extrinsic noise [11]. In this case, the

transcriptional response takes the form

$$\mathbf{z}g(\mathbf{x}) \quad (4.27)$$

where  $\mathbf{z}$  represents the number of copies of the plasmid. Alternatively  $\mathbf{z}$  could represent the number of molecules of the RNA polymerase, in which case, the transcriptional response for an auto-regulatory gene network with negative feedback would be

$$g(\mathbf{x}, \mathbf{z}) = \frac{k_0 \mathbf{z}}{1 + k_1 \mathbf{z} + k_2 \mathbf{x}} \quad (4.28)$$

where  $k_0$ ,  $k_1$  and  $k_2$  are positive constants [6].

We model the stochastic fluctuations in  $\mathbf{z}$  by a birth-death process. In particular, the probabilities of formation and degradation of  $\mathbf{z}$  in the infinitesimal time interval  $(t, t + dt]$  are given by

$$\Pr\{\mathbf{z}(t + dt) = z + \mathbf{N}_z \mid \mathbf{z}(t) = z\} = K_z dt \quad (4.29a)$$

$$\Pr\{\mathbf{z}(t + dt) = z - 1 \mid \mathbf{z}(t) = z\} = d_z z dt \quad (4.29b)$$

where  $K_z$  and  $d_z$  represent the production and degradation rate of  $\mathbf{z}$ , respectively, and  $\mathbf{N}_z$  is a random variable with mean  $N_z$  and variance  $V_z^2$ . In the sequel we refer to  $T_z = \ln(2)/d_z$  as the response time of the exogenous signal. Following steps similar to those outlined in Section 4.1, we can conclude from (4.10) and (4.11) that the steady-state average level and the coefficient of variation of  $\mathbf{z}$  are given by

$$\mathbf{z}^* = \frac{N_z K_z}{d_z} \quad (4.30)$$

and

$$CV_z = \sqrt{\frac{(N_z^2 + V_z^2 + N_z)}{2\mathbf{z}^* N_z}}, \quad (4.31)$$

respectively. The quantity  $CV_z$  represents the amount of noise that enters the auto-regulatory gene network through the exogenous signal  $\mathbf{z}$ . Assuming that the stochastic fluctuations in  $\mathbf{x}$  and  $\mathbf{z}$  around their respective means  $\mathbf{x}^*$  and  $\mathbf{z}^*$  are sufficiently small, we approximate the transcriptional response as

$$g(\mathbf{x}, \mathbf{z}) \approx g(\mathbf{x}^*, \mathbf{z}^*) + \frac{dg(\mathbf{x}, \mathbf{z}^*)}{d\mathbf{x}}|_{\mathbf{x}=\mathbf{x}^*}(\mathbf{x} - \mathbf{x}^*) + \frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}}|_{\mathbf{z}=\mathbf{z}^*}(\mathbf{z} - \mathbf{z}^*), \quad (4.32)$$

by ignoring quadratic and higher order terms in  $\mathbf{x} - \mathbf{x}^*$  and  $\mathbf{z} - \mathbf{z}^*$ . For simplicity of notation, in the sequel  $g(\mathbf{x})$  refers to  $g(\mathbf{x}, \mathbf{z}^*)$ , the transcriptional response when there is no noise in the exogenous signal. Details are presented in Appendix B.2, where we show that for this linearized transcriptional response,  $\mathbf{x}^*$  is the solution to (4.16) and the total protein noise  $CV_{tot}$  is given by

$$CV_{tot}^2 = CV_{int}^2 + CV_{ext}^2 \quad (4.33)$$

where  $CV_{int}$  is the previously computed intrinsic noise and

$$CV_{ext} = \frac{T_r}{T_{nr}} \sqrt{\frac{T_z}{T_z + T_r}} SCV_z, \quad S := \frac{\mathbf{z}^*}{g(\mathbf{x}^*, \mathbf{z}^*)} \frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}}|_{\mathbf{z}=\mathbf{z}^*}, \quad T_z := \frac{\ln(2)}{d_z} \quad (4.34)$$

represents the extrinsic noise in the protein population. Note that signals  $\mathbf{z}$  with small response times  $T_z$  result in smaller values of  $CV_{ext}$  because rapid fluctuations in the exogenous signal are “averaged out” by the dynamics of the gene network. Typically, only those exogenous signals that have response times much larger than the protein’s response time, contribute significantly to the extrinsic component of protein noise.

The extrinsic noise  $CV_{ext}$  is a monotonically increasing function of the protein response time  $T_r$ , which in turn is determined by the slope of the transcriptional response  $g(\mathbf{x})$  at  $\mathbf{x} = \mathbf{x}^*$  [see (4.18)]. This is in contrast to the intrinsic noise  $CV_{int}$  which is determined by the y-intercept of the tangent to the transcriptional response  $g(\mathbf{x})$  at  $\mathbf{x} = \mathbf{x}^*$  [see (4.20)]. Another important difference is that unlike  $CV_{int}$ ,  $CV_{ext}$  does not depend

on gene expression parameters such as the average number  $N_x$  and variance  $V_x^2$  of proteins produced per gene expression event.

We now contrast how rapidly intrinsic and extrinsic noises attenuate as the response time  $T_r$  is decreased. We first express  $CV_{ext}$  as a function of the extrinsic noise level  $CV_{ext-nr}$  that would be observed in the absence of feedback:

$$CV_{ext} = \frac{T_r}{T_{nr}} \sqrt{\frac{T_{nr} + T_z}{T_r + T_z}} CV_{ext-nr}, \quad (4.35)$$

where

$$CV_{ext-nr} = \sqrt{\frac{T_z}{T_z + T_{nr}}} SCV_z. \quad (4.36)$$

From (4.35), we conclude that the five fold decrease in the response time (i.e.,  $T_r \approx T_{nr}/5$ ) that we had observed in Figure 4.3 corresponds to a reduction of  $CV_{ext}$  by a factor of 3.9 compared to  $CV_{ext-nr}$  when  $T_z \approx T_{nr}$  or a reduction by a factor of 5 when  $T_z \gg T_{nr}$ . Recall from the previous section that a five fold decrease in the response time leads to a reduction of intrinsic noise level in the protein by a factor of only 2.2 (assuming that  $\mathbf{x}^*$  is kept fixed). This illustrates an important point: negative feedback is much more effective in reducing the extrinsic component of protein noise than its intrinsic component.

## 4.4 Auto-regulatory gene networks with negative feedback

In this section we quantify protein noise levels and limits of noise suppression for auto-regulatory gene networks involving a common form of transcriptional response

given by

$$g(\mathbf{x}) = g_0 \left( b + \frac{1-b}{1+(a\mathbf{x})^M} \right), \quad 0 \leq b < 1 \quad (4.37)$$

where  $M \geq 1$  denotes the Hill coefficient and  $g_0$  is the transcription rate when there is no feedback ( $a = 0$ ) [3, 58]. The constant  $b$  is chosen such that the product  $g_0 b$  represents the minimum level of transcription rate (also called the basal level of transcription rate) that is achieved when the number of protein molecules is very large ( $\mathbf{x} \rightarrow \infty$ ). Typically, the constant  $b$  is either zero or much smaller than one. The constant  $a$  is the feedback strength that essentially depends on the binding affinity of the protein to the promoter, with lower binding affinities corresponding to smaller values of  $a$ .

For the above transcriptional response we conclude from (4.16) that the equilibrium  $\mathbf{x}^*$  is the unique solution to

$$N_x g(\mathbf{x}^*) = N_x g_0 \left( b + \frac{1-b}{1+(a\mathbf{x}^*)^M} \right) = d_x \mathbf{x}^* \quad (4.38)$$

and monotonically decreases as we increase  $a$ . The response time  $T_r$  in (4.18) is given by

$$T_r = \frac{T_{nr}(1+(a\mathbf{x}^*)^M)(1+b(a\mathbf{x}^*)^M)}{1+[1+M-b(M-1)](a\mathbf{x}^*)^M+b(a\mathbf{x}^*)^{2M}}, \quad T_{nr} = \frac{\ln(2)}{d_x} \quad (4.39)$$

which starts by decreasing as we increase the feedback strength  $a$  from an initial value of  $a = 0$ . It achieves a minimum value of

$$T_{rmin} = T_{nr} \frac{1+\sqrt{b}}{1+M-\sqrt{b}(M-1)} \quad (4.40)$$

for

$$a = a_{Tmin} = \frac{b^{-\frac{M+1}{2M}} d_x}{g_0 N_x} \quad (4.41)$$

and then increases as we increase  $a$  beyond  $a_{Tmin}$ . Note that when  $b = 0$ , we have  $a_{Tmin} = \infty$  and therefore  $T_r$  always decreases as we increase the feedback strength  $a$  with the asymptote

$$\lim_{a \rightarrow \infty} T_r = \frac{T_{nr}}{M+1}. \quad (4.42)$$

In the sections below we investigate how the different components of the noise and the total noise in the protein numbers change as the feedback strength  $a$  varies.

#### 4.4.1 Suppression of intrinsic noise in the protein

We first investigate the intrinsic component of noise given by (4.21) for this specific transcriptional response. Substituting (4.37) in (4.21), and using (4.38) and (4.39), we conclude that the intrinsic noise  $CV_{int}$  in the protein is given by

$$CV_{int} = \sqrt{\frac{T_r}{T_{nr}} \frac{1 + (a\mathbf{x}^*)^M}{1 + b(a\mathbf{x}^*)^M}} CV_{int-nr} \quad (4.43a)$$

$$= \sqrt{\frac{[1 + (a\mathbf{x}^*)^M]^2}{1 + [1 + M - b(M-1)](a\mathbf{x}^*)^M + b(a\mathbf{x}^*)^{2M}}} CV_{int-nr} \quad (4.43b)$$

where

$$CV_{int-nr} = \sqrt{\frac{d_x(N_x^2 + V_x^2 + N_x)}{2g_0N_x^2}} \quad (4.44)$$

is the intrinsic noise in the protein when there is no feedback (i.e.,  $a = 0$ ). Our goal is to understand how  $CV_{int}^2$  varies with the Hill coefficient  $M$  and the feedback strength  $a$ . Straightforward calculus shows that the above intrinsic noise is smallest when the feedback strength is equal to

$$a_{int-min} = \frac{d_x}{N_x g_0} \frac{2M}{M+1+b(M-1)} \left( \frac{M-1}{M+1} \right)^{\frac{1}{M}} \quad (4.45)$$

and the corresponding minimum intrinsic noise  $CV_{int-min}$  is given by

$$CV_{int-min} = \sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} CV_{int-nr} \leq CV_{int-nr}. \quad (4.46)$$

When  $M = 1$ , then  $a_{int-min} = 0$  and  $CV_{int-min} = CV_{int-nr}$ , i.e., the intrinsic noise level is minimum when there is no feedback. In this particular case, increasing  $a$  causes  $CV_{int}$  to monotonically increase (see Figure 4.5). This happens because as we increase  $a$  from zero, both  $T_r$  and  $\mathbf{x}^*$  decrease in (4.21). However, as  $\mathbf{x}^*$  decreases at a faster rate than  $T_r$ , their ratio  $T_r/\mathbf{x}^*$  increases, and hence, the intrinsic noise increases as we increase the feedback strength  $a$ . When  $M > 1$ , this behavior changes and the intrinsic noise first decreases when we increase  $a$  starting from zero and achieves a minimum at some optimal value  $a = a_{int-min} > 0$ . Increasing  $a$  beyond  $a_{int-min}$  causes an increase in the intrinsic noise level (see Figure 4.5).

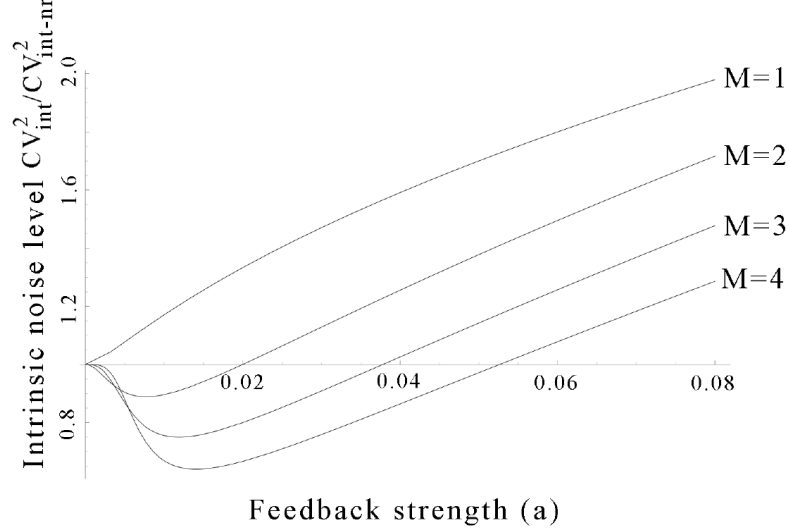


Figure 4.5. Intrinsic noise  $CV_{int}$  in the protein as a function of the feedback strength  $a$  and Hill coefficient  $M$ .  $CV_{int}$  is normalized by  $CV_{int-nr}$ , the intrinsic noise in the protein when there is no feedback. Other parameters taken as  $g_0 = 1$ ,  $b = 0$ ,  $N_x = 1$ ,  $V_x = 0$  and  $d_x = 0.01$ .

From (4.46), we conclude that the quantity

$$\frac{CV_{int-min}}{CV_{int-nr}} = \sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} \quad (4.47)$$

represents the highest suppression of intrinsic noise in the protein from  $CV_{int-nr}$  that can be achieved with the transcriptional response given by (4.37). Notice that this limit decreases as we decrease the basal level of transcription (i.e., decrease  $b$ ) and achieves a minimum at  $b = 0$ . For the (common) case  $b \ll 1$ , this limit of intrinsic noise suppression is simply given by

$$\frac{CV_{int-min}}{CV_{int-nr}} = \frac{\sqrt{4M}}{M+1} \quad (4.48)$$

and is completely determined by the Hill coefficient  $M$ , with larger values of  $M$  causing more reduction in the protein intrinsic noise. This is consistent with results in the literature, which show that a large Hill coefficient is more effective in reducing stochastic fluctuations in the protein [58, 35, 51]. For example, when  $b = 0$ , and  $M = 2$  there can be at most a  $1 - \sqrt{4M}/(M+1) = 5.7\%$  reduction in intrinsic noise from  $CV_{int-nr}$ , whereas for  $M = 4$  we can have a 20% reduction.

In summary, depending on the Hill coefficient, the protein intrinsic noise levels can either monotonically increase or exhibit a U-shaped curve as the feedback strength is increased. Moreover, large Hill coefficients are much more effective in reducing noise. The limit of noise suppressions computed above are good approximations when the mRNA half life is much smaller than the protein's half life. In Appendix B.7 we consider deviations from this case and show that when the mRNA half life is not small compared to the protein's half life, the limit of noise suppression is slightly smaller than what is predicted by (4.47) and (4.48).



#### 4.4.2 Suppression of extrinsic noise in the protein

We now investigate the extrinsic component of protein noise  $CV_{ext}$ . As  $CV_{ext}$  is a monotonically increasing function of the response time  $T_r$  [see equation (4.34)], it will be minimum when the response time is the smallest. Recall from (4.40), (5.9) that the response time  $T_r$  achieves its minimum value  $T_{rmin}$  when the feedback strength is equal to  $a = a_{Tmin}$ . Hence we conclude that the level of extrinsic noise is minimum when the feedback strength is equal to

$$a_{ext-min} = \frac{b^{-\frac{M+1}{2M}} d_x}{g_0 N_x}, \quad (4.49)$$

and from (4.35), this minimum level  $CV_{ext-min}$  is given by

$$\frac{CV_{ext-min}}{CV_{ext-nr}} = \sqrt{\frac{T_{nr} + T_z}{T_{rmin} + T_z} \frac{T_{rmin}}{T_{nr}}}, \quad T_{rmin} = \frac{1 + \sqrt{b}}{1 + M - \sqrt{b}(M-1)} T_{nr}. \quad (4.50)$$

The above expression provides the limit of extrinsic noise suppression and reduces to

$$\frac{CV_{ext-min}}{CV_{ext-nr}} = \begin{cases} \sqrt{\frac{T_{nr} + T_z}{[T_{nr} + T_z(M+1)](M+1)}} & \text{when } b \ll 1 \\ \frac{1}{M+1} & \text{when } b \ll 1 \text{ and } T_z \gg T_{nr}. \end{cases} \quad (4.51)$$

As we increase  $M$  these limits decrease at a much faster rate than the limit of intrinsic noise suppression for the same value of  $b$  [compare with right-hand-side of (4.48)]. For example, when  $T_z \approx T_{nr}$  and  $b \ll 1$ , for  $M = 2$  we have a maximum reduction in extrinsic noise of  $1 - \sqrt{2/[(M+1)(M+2)]} \approx 42\%$  whereas for  $M = 4$  we have a reduction of 74%. These reductions are much larger than the maximum reductions of 5.7% and 20% in the protein intrinsic noise level for the same values of  $M$  and  $b$  (see Section 4.4.1). This reinforces the earlier point that negative feedback is much more efficient in reducing the extrinsic component of the noise than its intrinsic component.

#### 4.4.3 Suppression of total noise in the protein

Finally, we investigate how the total noise in the protein population varies with the feedback strength. As derived in Section 4.3, the total protein noise level is given by

$$CV_{tot}^2 = CV_{int}^2 + CV_{ext}^2, \quad (4.52)$$

which using (4.34), (4.39) and (4.43) can be written as

$$CV_{tot}^2 = CV_{int-nr}^2 \frac{T_r}{T_{nr}} \frac{1 + (a\mathbf{x}^*)^M}{1 + b(a\mathbf{x}^*)^M} + S^2 CV_z^2 \left( \frac{T_r}{T_{nr}} \right)^2 \frac{T_z}{T_z + T_r} \quad (4.53a)$$

$$T_r = \frac{T_{nr}(1 + (a\mathbf{x}^*)^M)(1 + b(a\mathbf{x}^*)^M)}{1 + [1 + M - b(M - 1)](a\mathbf{x}^*)^M + b(a\mathbf{x}^*)^{2M}}. \quad (4.53b)$$

Now, for all  $M \geq 1$  and  $CV_z > 0$  we have that

$$\frac{dCV_{tot}^2}{da} \Big|_{a=0} = -(1 - b) \left[ (CV_{int-nr}^2)(M - 1) + CV_z^2 MT_z \frac{(2T_z + T_{nr})}{(T_z + T_{nr})^2} \right] < 0, \quad (4.54)$$

which means that in the presence of extrinsic noise, the total protein noise level will always decrease as we increase the feedback strength from  $a = 0$ , irrespective of the value of the Hill coefficient, but eventually will start to increase for sufficiently large values of  $a$  past an optimal feedback strength  $a_{min}$ . In summary, in the presence of extrinsic noise, the total noise in the protein is always minimized at some optimal feedback strength and decreasing or increasing feedback strength away from this optima will always causes an increase in the noise level. This point is shown in Figure 4.6 which plots  $CV_{tot}/CV_{tot-nr}$  as a function of  $a$  when the Hill coefficient is one, where

$$CV_{tot-nr}^2 = CV_{int-nr}^2 + S^2 CV_z^2 \frac{T_z}{T_z + T_{nr}}, \quad (4.55)$$

represents the protein noise level when there is no feedback. We can see that in the absence of extrinsic noise ( $CV_z = 0$ ),  $CV_{tot}/CV_{tot-nr}$  monotonically increases as the feedback strength is increased. However, in the presence of extrinsic noise, it follows a U-shaped profile and is minimized at some  $a = a_{min} > 0$ .

As shown in Figure 4.7, when the Hill coefficient is larger than one ( $M > 1$ ), then even in the absence of any extrinsic noise ( $CV_z = 0$ ), the protein noise level will show a U-shaped profile as the feedback strength is altered (see Section 4.4.1). In particular, for  $CV_z = 0$ , we conclude from (4.47) that the minimum value of  $CV_{tot}/CV_{tot-nr}$ , i.e., the limit of noise suppression is given by

$$\frac{CV_{tot-min}}{CV_{tot-nr}} = \sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} \quad (4.56)$$

and is attained when the feedback strength is equal to

$$a_{min} = \frac{d_x}{N_x g_0} \frac{2M}{M+1+b(M-1)} \left( \frac{M-1}{M+1} \right)^{\frac{1}{M}}. \quad (4.57)$$

As shown in Figure 4.6 (for  $M = 1$ ) and Figure 4.7 (for  $M = 2$ ), when we now increase  $CV_z$  away from zero, this limit of noise suppression decreases and is much lower than what is predicted by (4.56). On the other hand the optimal feedback strength  $a_{min}$  at which the protein noise is minimum, increases and is much higher than (4.57). As we further increase the noise  $CV_z$  in the exogenous signal, both  $CV_{tot-min}/CV_{tot-nr}$  and  $a_{min}$  approach (4.50) and (4.49), respectively, which correspond to the scenario where extrinsic noise dominates the total noise in protein numbers.

In Appendix B.3, we provide formulas that predict both the minimum level of noise  $CV_{tot-min}$  and the optimal feedback strength when both intrinsic and extrinsic noise are present but neither one dominates the total noise in the protein population. As we will see shortly, an important application of these formulas is that they can be used to estimate the noise in the exogenous signal from the experimentally obtained value of  $CV_{tot-min}$ , *without directly measuring the exogenous signal*.

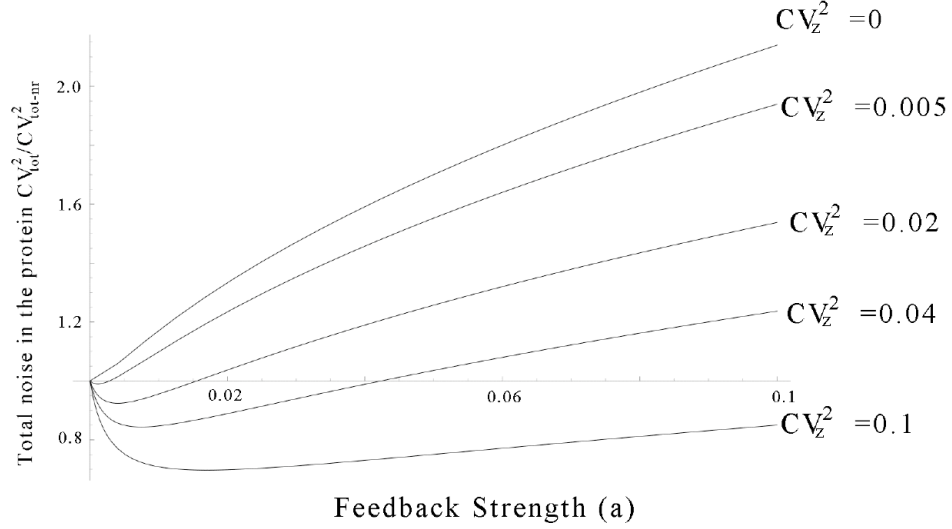


Figure 4.6. Total noise  $CV_{tot}$  as a function of the feedback strength  $a$  when the Hill coefficient is one ( $M = 1$ ) for different values of noise  $CV_z$  in the exogenous signal.  $CV_{tot}$  is normalized by  $CV_{tot-nr}$ , the total noise in the protein when there is no feedback. Other parameters are taken as  $g_0 = 1$ ,  $N_x = 4$ ,  $V_x^2 = N_x^2 - N_x$ ,  $b = 0$ ,  $S = 1$ ,  $d_x = 0.04$ . The response time  $T_z$  is assumed to be much larger than  $T_{nr}$ .

## 4.5 Experimental verification

We now validate our theoretical results with recent experimental measurements of protein noise levels that were obtained as the feedback strength was changed via experimental manipulation.

In [11], a synthetic auto-regulatory gene network shown in Figure 4.8 is constructed, where the protein inhibits its own transcription. The feedback strength is altered by adding a compound aTc that binds to the protein and the resulting complex has a significantly smaller binding affinity to the promoter. As the feedback strength is directly related to the binding affinity of the protein to its promoter, increasing the concentration of aTc corresponds to decreasing the feedback strength  $a$ . The gene is encoded on a low-copy plasmid with high variability in plasmid population contribut-

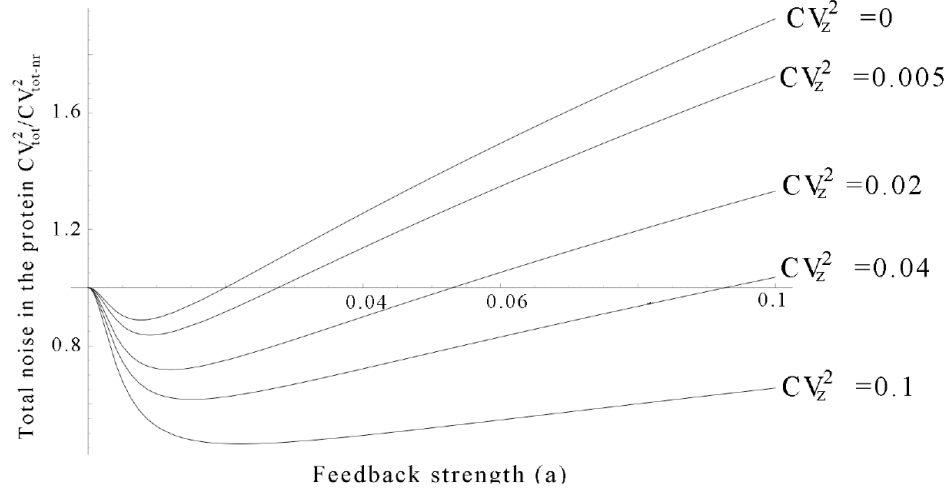


Figure 4.7. Total noise  $CV_{tot}$  as a function of the feedback strength  $a$  when the Hill coefficient is two ( $M = 2$ ) for different values of noise  $CV_z$  in the exogenous signal.  $CV_{tot}$  is normalized by  $CV_{tot-nr}$ , the total noise in the protein when there is no feedback. Other parameters are taken as  $g_0 = 1$ ,  $N_x = 4$ ,  $V_x^2 = N_x^2 - N_x$ ,  $b = 0$ ,  $S = 1$ ,  $d_x = 0.04$ . The response time  $T_z$  is assumed to be much larger than  $T_{nr}$ .

ing to large levels of extrinsic noise in the protein population. Based on our theoretical analysis, the protein noise level should show a U-shaped profile as the feedback strength is changed. In particular, at low values of  $a$  (i.e., high levels of  $aTc$ ), increasing  $a$  (i.e., decreasing  $aTc$ ) should lead to a decrease in protein noise levels. However, at high values of  $a$  (i.e., low levels of  $aTc$ ), increasing  $a$  (i.e., decreasing  $aTc$ ) should increase the protein noise levels. Such a U-shaped profile is indeed experimentally observed and the protein noise level is minimized at an optimal level of feedback strength (see bottom left plot of Figure 4 in [11]).

In [11], the results from detailed stochastic simulations of the auto-regulatory gene network are also reported. The authors observe in simulation that both in the absence of any extrinsic noise or when the extrinsic noise from only the enzyme RNA polymerase

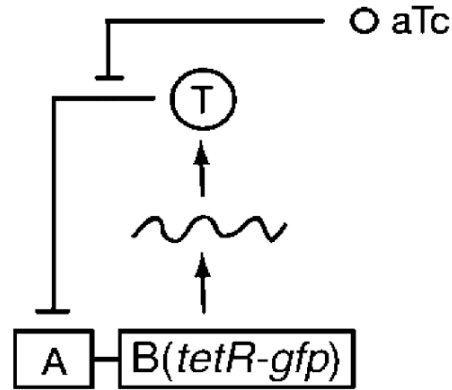


Figure 4.8. A synthetic auto-regulatory gene network where the feedback strength is manipulated by adding aTc. This figure is taken from [11].

is included, instead of seeing a U-shaped profile, the protein noise level monotonically increased as the feedback strength is increased (i.e., aTc concentration is decreased). Our theoretical results fully explain this phenomenon: Since in this synthetic gene network the Hill coefficient is one ( $M = 1$ ), our analysis in Section 4.4.1 shows that the intrinsic noise level will always increase when the feedback strength is increased. As the extrinsic noise associated with fluctuations in RNA polymerase numbers is very small (we calculate  $CV_{RNA\ polymerase} \approx 0.02$  using (4.31) and the reaction rates provided in Table I of [11]), in both the above cases the protein noise is dominated by the intrinsic noise, which always increases with the feedback strength, and hence, no U-shaped profile should be observed.

As mentioned earlier, our results also allow us to predict the level of noise in the exogenous signal that drives the synthetic auto-regulated gene network. Hypothesizing that the source of extrinsic noise is the plasmid population, and using the experimentally obtained minimal protein noise level of approximately 0.4, we estimate using the

formulas derived here that

$$CV_{plasmid} \approx 0.64 \quad (4.58)$$

(see Appendix B.4). Independent measurements of plasmid noise (using (4.31) and the reaction rates provided in Table I of [11]) show that  $CV_{plasmid}$  is equal to 0.51 which is just slightly smaller than our estimate in (4.58). This indicates that variability in plasmid numbers is indeed the major source of extrinsic noise in the protein population. The fact that the estimate in (4.58) is larger than the actual plasmid noise suggests that variability in other cellular components also make (minor) contributions to the extrinsic noise.

In summary, the experimental results of [11] provide an experimental verification of our theoretical predictions. They also show that measuring changes in the protein noise level as a function of the feedback strength can be used to determine the level of noise in the exogenous signal. Alternatively, our results can be used to confirm hypothesized sources for extrinsic noise.

## 4.6 Discussion

Auto-regulatory gene networks where the protein inhibits/activates its own transcription are common motifs occurring within living cell. These networks are characterized by their transcriptional response  $g(\mathbf{x})$  which provides information on how the transcription rate of the gene varies as a function of the number of protein molecules  $\mathbf{x}$  present in the cell.

### 4.6.1 Noise dependence on the shape of the transcriptional response

We developed a full understanding of how the protein noise levels are related to the functional form of the transcriptional response. Using a linear approximation for  $g(\mathbf{x})$ , we showed that the extrinsic noise levels are determined by the slope  $g'(\mathbf{x}^*)$  of the transcriptional response at  $\mathbf{x}^*$ , with more negative values of the slope (i.e., more stable equilibria  $\mathbf{x}^*$ ) leading to smaller levels of extrinsic noise. On the other hand, the intrinsic noise levels are determined by  $I = g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)$  which is the y-intercept of the tangent to the transcriptional response at  $\mathbf{x} = \mathbf{x}^*$  (as shown in Figure 4.2), and larger values of  $I$  lead to smaller levels of intrinsic noise. Consequently, given two hypothetical transcriptional responses  $g_1(\mathbf{x}) = 1$  and  $g_2(\mathbf{x}) = 1 - \mathbf{x}/2$ , the response  $g_2(\mathbf{x})$  will give lower levels of extrinsic noise. However, since both transcriptional responses have the same intercept  $I$  equal to one (see Figure 4.9), they both yield the same level of intrinsic noise in the protein population.

We also considered deviations from a linear transcriptional response and showed that concave responses have better noise suppression properties than convex responses.

### 4.6.2 Intrinsic v.s. extrinsic noise

Analytical formulas that relate the noise levels to the response time of the protein show key differences between extrinsic and intrinsic noise: as one decreases the protein response time  $T_r$  through feedback, the levels of extrinsic noise decreases much more than those of intrinsic noise. This leads to an important conclusion that negative feedback is much more effective in reducing the extrinsic component of protein noise than its intrinsic component, which is consistent with other theoretical and experimental studies [37, 20, 47, 55]. Another difference is that unlike intrinsic noise,



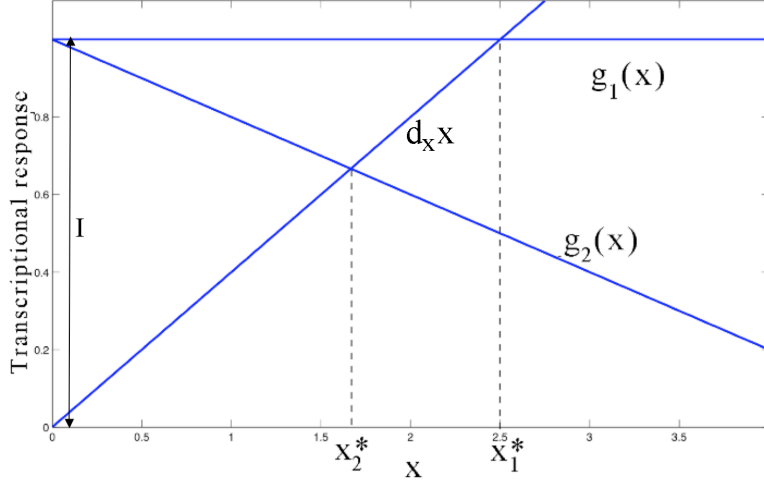


Figure 4.9. Two different transcriptional responses  $g_1(\mathbf{x})$  and  $g_2(\mathbf{x})$  that lead to the same intrinsic noise in the protein.  $\mathbf{x}_1^*$  and  $\mathbf{x}_2^*$  represents the steady-state average number of protein molecules when the transcriptional response is given by  $g_1(\mathbf{x})$  and  $g_2(\mathbf{x})$ , respectively.

the extrinsic noise is independent of the average burst size  $N_x$  and the variance  $V_x^2$  in the number of protein molecules produced in each transcription event. Recent work [13] has suggested that many genes operate with very low values of  $N_x$ , which could be an adaptation to reduce the intrinsic noise but that, according to our results, appears to have no effect on the extrinsic noise levels.

### 4.6.3 U-shaped protein noise profile

We investigated how protein noise levels change as we vary the feedback strength for a biologically meaningful class of auto-regulatory gene networks with negative feedback characterized by the transcriptional response

$$g(\mathbf{x}) = g_0 \left( b + \frac{1-b}{1+(a\mathbf{x})^M} \right), \quad 0 \leq b < 1. \quad (4.59)$$

Our main result shows that the total noise level in the protein population is minimized at an optimal level of feedback strength. Recall from Section 4.4 that increasing the feedback strength causes a decrease in the average number of protein molecules, which results in an increase in the intrinsic noise level. On the other hand increasing the feedback strength causes the protein response time to decrease which attenuates both the intrinsic and extrinsic noise. The net result of these two opposing effects is a U-shaped profile, where increasing feedback strength first causes the noise level to decrease and then increase. This U-shaped profile was shown to be in good agreement with experimental data for a synthetic auto-regulatory gene network. We also identified a scenario where noise is minimum when there is no feedback and any amount of negative feedback will always increase the noise: the case where intrinsic noise dominates the total noise in the protein population and the Hill coefficient is close to one. This explained an observation in [11] that when the source of the extrinsic noise was removed, the U-shaped profile vanished, and instead, the noise level monotonically increased with the feedback strength. However, for synthetic gene networks characterized by a Hill coefficient much larger than one, our theoretical results predicts that even in the absence of extrinsic noise a U-shaped profile should be observed. This remains to be experimentally verified.

#### **4.6.4 Limit of noise suppression**

We characterized the smallest level of noise that is inherent to this type of auto-regulation. This was done through the limit of noise suppression which is defined to be the ratio of the minimum possible noise with feedback to the protein noise level when there is no feedback (i.e.,  $a = 0$ ) and corresponds to the depth of the U-shape profile in Figure 4.6 and Figure 4.7. For auto-regulatory networks with a small basal

level of transcription (i.e.,  $b \approx 0$ ) this limit is given by

$$\frac{\sqrt{4M}}{M+1} \quad (4.60)$$

when the intrinsic noise dominates the total noise in the protein [see (4.48)]. However, as the amount of extrinsic noise increases this limit decreases and asymptotically approaches

$$\sqrt{\frac{T_{nr} + T_z}{[T_{nr} + T_z(M+1)](M+1)}}, \quad (4.61)$$

which corresponds to the situation where extrinsic noise completely dominates the total noise in the protein [see (4.50)].

We also showed that the optimal level of feedback strength, for which the protein noise level is minimum, monotonically increases with increasing levels of extrinsic noise in the protein population. Thus if negative feedback loops indeed function to minimize protein noise, then networks with larger contributions from extrinsic noise compared to intrinsic noise, should operate at higher levels of feedback strength, i.e., higher binding affinities between the protein and its promoter.

The above results can be used to quantify the level of extrinsic noise in the protein population. This is useful for synthetic and natural auto-regulatory gene networks where the feedback strength can be manipulated. As illustrated in Section 4.5, noise in the exogenous signal driving the extrinsic noise can be estimated from the minimum possible protein noise. Matching these estimates with independent measurements of noise in the exogenous signal can be used to confirm hypothesis that a particular noise source is the major contributor of extrinsic noise to the protein population.

### 4.6.5 Positive feedback loops

Our analysis can also be used for auto-regulatory gene networks with positive feedback. These networks are characterized by a transcriptional response  $g(\mathbf{x})$  as in (4.59), but with a constant  $b$  larger than one. A similar analysis reveals that instead of being minimized, the protein noise levels are maximized at the optimal level of feedback strength. Thus protein noise levels follow an inverted U-shape profile as the feedback strength is increased. Moreover, for positive feedback loops characterized by a Hill coefficient of one (such as the Tat protein in the HIV gene network [63, 62]), and when intrinsic noise dominates the total noise, the noise level monotonically decrease with increasing feedback strength. Such differences in noise profiles can again be exploited to determine the level of extrinsic noise in these gene networks.

In summary, we have developed results relating the noise levels to the feedback strength in auto-regulatory gene networks. We have shown that for negative feedback loops, protein noise levels are always minimized at an optimal level of feedback strength. The noise resulting from these optimal levels of feedback characterizes the smallest level of noise that can be achieved in these networks through the use of negative feedback. Our results have implications for the design of synthetic auto-regulatory gene networks with minimal protein noise. They also raise the question of whether these widely occurring auto-regulatory gene networks (for example, over 40% of known *Escherichia coli* transcription factors negatively regulate their own transcription [45]) have naturally evolved to operate at the optimal feedback strength. If these networks have indeed evolved to operate at the optimal point then any mutation or experimental manipulation that changes the feedback strength should always cause the protein noise levels to increase.

## 4.7 Future work

We derived analytical formulas for the protein noise level using the linear noise approximation which involves linearizing the transcriptional response  $g(\mathbf{x}, \mathbf{z})$  about the means  $\mathbf{x}^*$  and  $\mathbf{z}^*$ . As pointed out this is a valid approximation as long as the protein noise level is sufficiently small. One direction of future work is to extend this analysis to situations where the protein noise level is not small, and understand how the protein noise level is related to both the functional form of  $g(\mathbf{x}, \mathbf{z})$  and how the exogenous signal  $\mathbf{z}$  enters the function  $g$ .

We developed a qualitative understanding of how protein noise levels change with the feedback strength when the transcriptional response  $g(\mathbf{x}, \mathbf{z})$  is given by the Hill function. This typically correspond to a situation where the protein molecules combines with other copies of the protein to form a functional multimer which then binds to the promoter switching the gene off. The Hill function would be a valid approximation only when the dynamics of protein multimerization and binding/unbinding of the multimer to the promoter is sufficiently fast. A direction of future work is to relax this assumption and explicitly model both the above processes in the auto-regulatory gene network. For example, protein multimerization can be modeled as a linear cascade where proteins combine to form dimers, which then combines with proteins to form trimers, and so on. Our goal would be to understand how protein noise level changes as the feedback strength is varied when the dynamics of the protein multimerization and binding/unbinding of the multimer to the promoter is not fast compared to the dynamics of protein production and degradation.

Another direction of future work is to follow up the predictions made in Section 4.2.2 about the lambda repressor gene network. In Section 4.2.2 we proposed

that the negative feedback reduces stochastic fluctuations in protein numbers and allows the virus to have a more robust lysogeny. This is contrary to the current view that negative feedback allows the virus to have less number of lambda repressors, which help it respond faster when a signal to lyse the cell is activated [42]. Using more detailed stochastic models of the lambda repressor gene network (see [4]), our goal would be to analyze the role of the negative feedback.

## **Chapter 5**

# **Evolution of auto-regulatory gene networks**

In this chapter we investigate under what conditions an auto-regulatory negative feedback mechanism can evolve from a primitive gene network with no auto-regulation. In the course of evolution, negative feedbacks would be introduced in the primitive network through random mutations in the protein and the promoter region of its gene. However, these mutations would only persist if the resulting feedback increases the fitness of the corresponding cell line.

We consider a simple form for the fitness function that is given by the probability of maintaining protein numbers above a critical level. This fitness function is appropriate for various essential proteins whose populations have to be maintained above a threshold for normal cellular functioning. Stochastic fluctuations in the protein population that drive their numbers below this threshold are assumed to compromise the cell's viability. This type of threshold like responses have been reported for various prokaryotic

transcription factors [52]. The above fitness functions is also appropriate for genes that can have a stable ON and OFF state corresponding to different environmental inputs (for example, the lambda repressor gene in the lambda phage gene network [42] and the Gal80 gene in the galactose signaling network [1]). For such genes, minimizing random stochastic transitions from the ON to the OFF state correspond to maximizing the probability of having protein numbers above a critical threshold.

The fitness function considered here results in a nontrivial tradeoff for the evolution of a negative feedback. Any mutation that introduces negative feedback decreases the average number of protein molecules, which tends to decrease the probability of having protein numbers larger than the threshold and hence decreases fitness. However, the negative feedback will also reduce fluctuations in protein numbers about the mean, which tends to increase this same probability. If the net change in fitness is positive, then the cell line with the negative feedback persists. Moreover, subsequent additional mutations to other elements of the auto-regulatory network can even bring the average number of protein molecules back to their original level. However, if the net change in fitness is negative, then the cell line dies and evolution of negative feedback is not possible.

In this chapter we perform a systematic analysis of the above fitness function and provide conditions under which a negative feedback is evolvable. In particular, this thesis shows that a negative feedback is more evolvable if the source of noise in the protein population is extrinsic and not intrinsic noise. For example, a negative feedback with a Hill coefficient of one can never evolve if fluctuations in protein numbers are intrinsic, but this feedback can evolve if there is sufficiently large extrinsic noise in the protein population. We also make other predictions that a negative feedback is more evolvable when the Hill coefficient of the feedback is large, the response time of



the exogenous signal is large, and the critical threshold level is small.

## 5.1 Fitness function

We consider the situation where the protein is only active when its population is above a critical threshold  $\mathbf{T}$ . Fluctuations in the population that causes the protein numbers to drift below this threshold inactivate the protein and it can no longer perform its biological function. Towards that end, our goal is to maximize the probability that the number of protein molecules  $\mathbf{x}$  is higher than the critical threshold  $\mathbf{T}$ . Assuming that the fluctuation in protein numbers about its mean is sufficiently small, and that the steady-state distribution of  $\mathbf{x}$  is approximately gaussian distribution, we have that

$$\text{Probability}\{\mathbf{T} < \mathbf{x}\} = \frac{1}{\sigma^* \sqrt{2\pi}} \int_{x=\mathbf{T}}^{\infty} \exp\left(-\frac{(x - \mathbf{x}^*)^2}{2\sigma^{*2}}\right) dx \quad (5.1)$$

where  $\mathbf{x}^*$  represents the steady-state average number of protein molecules and  $\sigma^*$  represents the steady-state standard deviation about  $\mathbf{x}^*$ . Since this probability should be high, the threshold  $\mathbf{T}$  should be smaller than the average number of proteins  $\mathbf{x}^*$ . Using the coordinate transformation

$$y = \frac{x - \mathbf{x}^*}{\sqrt{2}\sigma^*}, \quad (5.2)$$

we re-write the probability (5.1) as

$$\text{Probability}\{\mathbf{T} < \mathbf{x}\} = \frac{1}{\sigma^* \sqrt{2\pi}} \int_{x=\mathbf{T}}^{\infty} \exp\left(-\frac{(x - \mathbf{x}^*)^2}{2\sigma^{*2}}\right) dx \quad (5.3a)$$

$$= \frac{1}{2} \left( 1 + \text{erf}\left(\frac{\mathbf{x}^* - \mathbf{T}}{\sqrt{2}\sigma^*}\right) \right) \quad (5.3b)$$

where erf refers to the error function and is given by

$$\text{erf}\left(\frac{\mathbf{x}^* - \mathbf{T}}{\sqrt{2}\sigma^*}\right) = \int_{y=0}^{\frac{\mathbf{x}^* - \mathbf{T}}{\sqrt{2}\sigma^*}} \exp(-y^2) dy. \quad (5.4)$$

As the error function defined above is a monotonically increasing function, we have from (5.3b) that the Probability $\{\mathbf{T} < \mathbf{x}\}$  is a monotonically increasing function of  $(\mathbf{x}^* - \mathbf{T})/\sigma^*$ . Towards that end, we define the fitness of the cell as

$$\text{Fitness} = \frac{\mathbf{x}^* - \mathbf{T}}{\sigma^*} \quad (5.5)$$

and increasing (decreasing) this fitness corresponds to increasing (decreasing) the probability that the protein population is above the critical threshold. In order to study the effect of feedback in the cell fitness, it will be convenient to work with a *normalized fitness*  $\mathbf{F}$  defined by

$$\mathbf{F} = \frac{\text{Fitness with feedback}}{\text{Fitness without feedback}} \quad (5.6)$$

which using (5.5) can be written as

$$\mathbf{F} = \left( \frac{\sigma_{nr}^*}{\sigma^*} \right) \left( \frac{\mathbf{x}^* - \mathbf{T}}{\mathbf{x}_{nr}^* - \mathbf{T}} \right). \quad (5.7)$$

where  $\mathbf{x}_{nr}^*$  and  $\sigma_{nr}^*$  correspond to the steady-state average number of molecules and standard deviation when there is no negative feedback. In the next section we analyze under what conditions the normalized fitness  $\mathbf{F}$  will increase to a value larger than 1 when an auto-regulation mechanism is introduced in gene expression.

## 5.2 Evolvability of the negative feedback

We consider the auto-regulatory gene network model introduced in the previous chapter where the transcriptional response is given by

$$g(\mathbf{x}) = g_0 \left( b + \frac{1-b}{1 + (a\mathbf{x})^M} \right), \quad 0 \leq b < 1, \quad (5.8)$$

where  $M \geq 1$  denotes the Hill coefficient,  $a$  characterizes the feedback strength, and  $g_0$  corresponds to the transcription rate when there is no feedback (i.e.,  $a = 0$ ). We recall that for the above transcriptional response, the average number of protein molecules  $\mathbf{x}^*$  is given by

$$N_x g_0 \left( b + \frac{1-b}{1 + (a\mathbf{x}^*)^M} \right) = d_x \mathbf{x}^* \quad (5.9)$$

[see (4.38)] and it monotonically decreases with increasing feedback strength. The steady-state coefficient of variation is given by

$$CV_{tot}^2 = CV_{int-nr}^2 \frac{T_r}{T_{nr}} \frac{1 + (a\mathbf{x}^*)^M}{1 + b(a\mathbf{x}^*)^M} + S^2 CV_z^2 \left( \frac{T_r}{T_{nr}} \right)^2 \frac{T_z}{T_z + T_r} \quad (5.10)$$

where,

$$\frac{T_r}{T_{nr}} = \frac{(1 + (a\mathbf{x}^*)^M)(1 + b(a\mathbf{x}^*)^M)}{1 + [1 + M - b(M-1)](a\mathbf{x}^*)^M + b(a\mathbf{x}^*)^{2M}}, \quad (5.11a)$$

$$CV_{int-nr} = \sqrt{\frac{d_x(N_x^2 + V_x^2 + N_x)}{2g_0N_x^2}} \quad (5.11b)$$

is the intrinsic noise in the protein when there is no feedback,  $T_z$  is the response time of the exogenous signal, and  $CV_z$  denotes the noise in the exogenous signal [see (4.53)].

We begin by expressing the normalized fitness  $\mathbf{F}$  as a function of the coefficient of variation of  $\mathbf{x}$ . To this effect, we re-write (5.7) as

$$\mathbf{F} = \left( \frac{CV_{tot-nr}}{CV_{tot}} \right) \left( \frac{1 - \gamma \frac{\mathbf{x}_{nr}^*}{\mathbf{x}^*}}{1 - \gamma} \right) \quad (5.12)$$

where  $CV_{tot-nr}$  denotes the coefficient of variation of the protein population when there is no feedback and is given by

$$CV_{tot-nr}^2 = CV_{int-nr}^2 + S^2 CV_z^2 \frac{T_z}{T_z + T_{nr}}, \quad (5.13)$$

and the constant  $\gamma$  is a *normalized threshold* defined by

$$\gamma = \frac{\mathbf{T}}{\mathbf{x}_{nr}^*} < 1. \quad (5.14)$$

Any mutation that now increases the feedback strength from zero decreases the average protein number  $\mathbf{x}^*$  from  $\mathbf{x}_{nr}^*$  which causes  $\mathbf{F}$  to decrease [see (5.12)]. On the other hand, this same mutation will introduce a negative feedback and decrease  $CV_{tot}$ , which causes  $\mathbf{F}$  to increase. If the net change in the fitness of the cell is positive then the negative feedback is said to be *evolvable*. Our goal is to analyze under what conditions and for what range of parameters is a negative feedback mechanism evolvable.

By computing the derivative of  $F$  with respect to the parameter  $a$ , and after straightforward but tedious computations one concludes that  $F$  will increase as  $a$  increases from zero if and only if

$$CV_{int-nr}^2 [M(1 - \gamma) - (1 + \gamma)] + \frac{S^2 CV_z^2}{(1 + T_{nr}/T_z)^2} [(2 + T_{nr}/T_z)M(1 - \gamma) - 2(1 + T_{nr}/T_z)\gamma] \quad (5.15)$$

is positive.

We now consider the limiting situations where intrinsic or extrinsic noise dominate the total noise in the protein population. When the extrinsic noise is negligible ( $CV_z \approx 0$ ) and therefore the intrinsic noise dominates, the normalized fitness  $F$  increases if and only if

$$M > M_{int} = \frac{1 + \gamma}{1 - \gamma} > 1, \quad 0 < \gamma < 1, \quad (5.16)$$

where  $M_{int}$  is a critical value of the Hill coefficient, above which the negative feedback is evolvable (when intrinsic noise dominates). This shows that when fluctuations in protein numbers are mostly intrinsic, i.e., the stochastic fluctuations are mostly due to random protein expression and degradation events, a negative feedback with a Hill coefficient of one ( $M = 1$ ) is not evolvable, irrespective of the value  $\gamma$  of the normalized threshold. In such a scenario the feedback can only evolve if the Hill coefficient is sufficiently large. As  $M_{int}$  is a monotonically increasing function of  $\gamma$ , we also conclude

that negative feedback is more likely to evolve if the normalized threshold  $\gamma$  is small. For example, when  $\gamma = 1/2$ , the feedback is evolvable only for  $M > M_{int} = 3$ , while for  $\gamma = 1/3$  it is evolvable for  $M > M_{int} = 2$ .

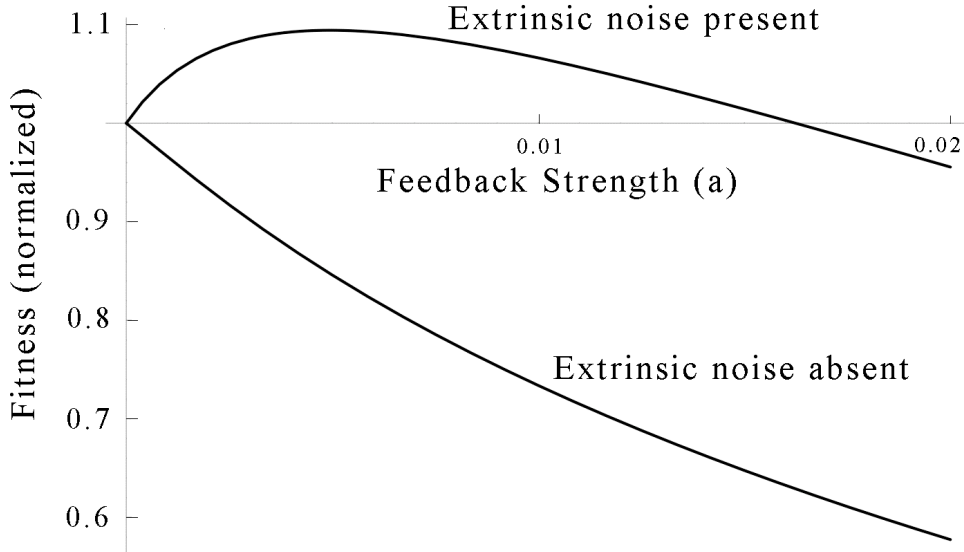


Figure 5.1. The normalized fitness  $\mathbf{F}$  as a function of the feedback strength  $a$  when  $M = 1$ . When extrinsic noise is absent the fitness decreases with increasing  $a$ , but in the presence of significant extrinsic noise, it increases. Other parameters taken as  $\gamma = 1/3$ ,  $\kappa = 0$ ,  $g_0 = 1$ ,  $N_x = 1$ ,  $V_x = 0$  and  $d_x = 0.01$ .

We next consider the other limiting case where extrinsic noise dominate the total noise in the protein population. In such a case we have from (5.15) that the negative feedback is evolvable if and only if

$$M > M_{ext} = \frac{2(1 + \kappa)\gamma}{(2 + \kappa)(1 - \gamma)} = \begin{cases} \frac{\gamma}{1 - \gamma} & \text{when } \kappa = \frac{T_{nr}}{T_z} \ll 1 \\ \frac{2\gamma}{1 - \gamma} & \text{when } \kappa = \frac{T_{nr}}{T_z} \gg 1 \end{cases}, \quad 0 < \gamma < 1, \quad (5.17)$$

where  $M_{ext}$  is the critical value of the Hill coefficient above which the negative feedback is evolvable (when extrinsic noise dominates). The quantity  $\kappa$  denotes the ratio

between the protein response time when there is no feedback and the response time of the exogenous signal driving the extrinsic noise.

Comparing (5.17) with (5.16) we see that  $M_{int} > M_{ext}$ , which implies that for Hill coefficients  $M$  in the range  $M_{ext} < M < M_{int}$ , negative feedback can only evolve if there is significant extrinsic noise in the protein population. In particular, when intrinsic noise dominates the total noise in the protein population, a negative feedback with  $M = 1$  cannot evolve irrespective of  $\gamma$ . However, when extrinsic noise dominates and  $\gamma$  is small enough such that  $M_{ext} < 1$ , then a negative feedback with  $M = 1$  is evolvable. This point is illustrated in Figure 5.2 which plots the normalized fitness  $\mathbf{F}$  as a function of the feedback strength  $a$  when  $M = 1$ .

Finally, we also point out that  $M_{ext}$  is a monotonically decreasing function of  $\kappa$  [see (5.17)]. This suggests that negative feedback is more likely to evolve when  $\kappa$  is small, i.e., the response time of the exogenous signal is large compared to that of the protein.

The results presented in this section are summarized in Figures 5.2 and 5.3. These figures show the evolvability of the negative feedback as a function of the Hill coefficient  $M$  and normalized threshold  $\gamma$  for  $\kappa \ll 1$  and  $\kappa \gg 1$ , respectively. In particular, the entire parameter space can be divided into three regions. The first region represents the case where a negative feedback is always evolvable irrespective of whether protein noise is intrinsic or extrinsic. This region corresponds to the parameter space where the Hill coefficient is large and the normalized threshold is small. The second region represents the parameter space where the negative feedback can only evolve if there is sufficiently large extrinsic noise in the protein population. As discussed above, feedback with  $M = 1$  corresponds to this region. Finally, the third region represents the parameter space where feedback is never evolvable and typically corresponds to

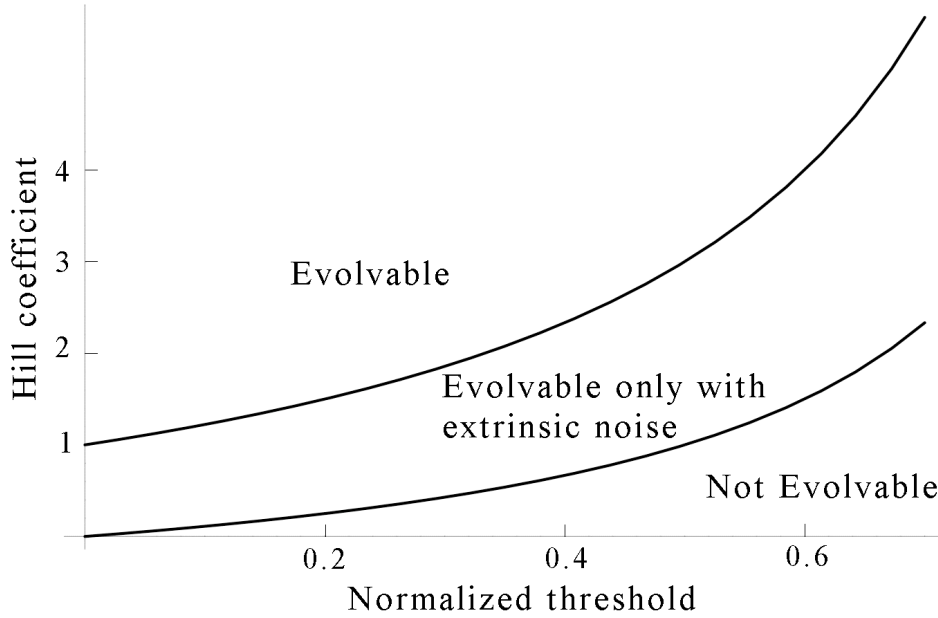


Figure 5.2. A plot of the evolvability of the negative feedback as a function of the Hill coefficient  $M$  and normalized threshold  $\gamma$  when  $\kappa \ll 1$ .

the situation where  $\gamma$  is large, i.e., the threshold  $\mathbf{T}$  is very near the average number of protein molecules. Comparing Figure 5.2 with Figure 5.3 we also see that the region that corresponds to evolvability only in the presence of extrinsic noise shrinks as  $\kappa$  is increased, i.e., the response time of the exogenous signal is decreased.

### 5.3 Discussion

Auto-regulatory gene networks in which a protein expressed by a gene inhibits its own transcription are common gene motifs within cells [3]. This chapter considered the evolution of such auto-regulatory gene networks from a primitive network with no auto-regulation. Using a fitness function motivated by a threshold like response

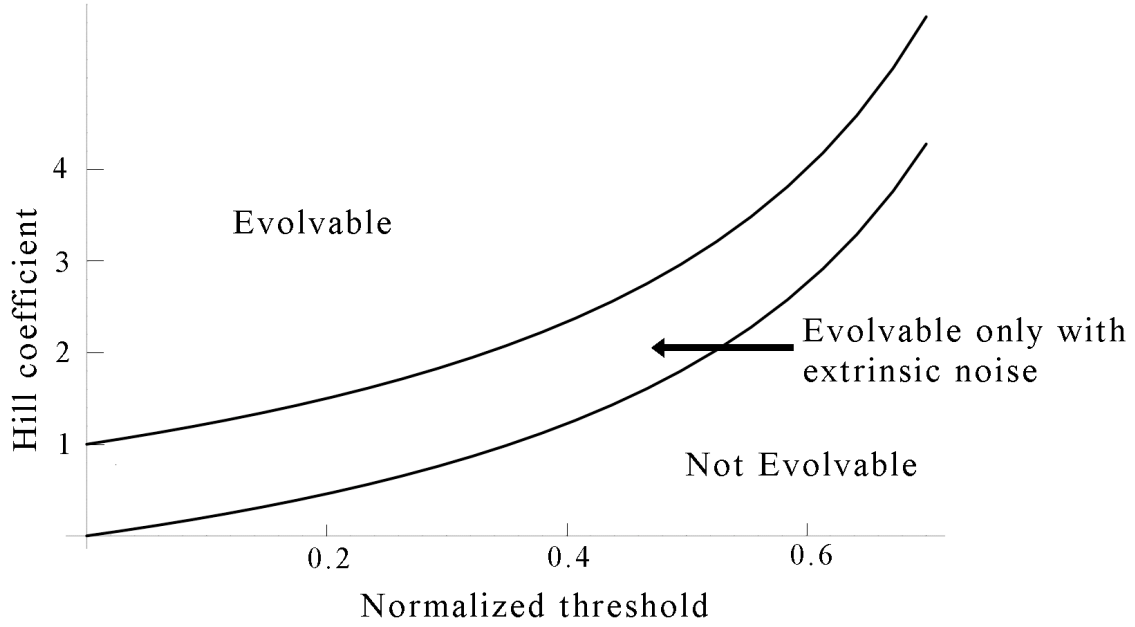


Figure 5.3. A plot of the evolvability of the negative feedback as a function of the Hill coefficient  $M$  and normalized threshold  $\gamma$  when  $\kappa \gg 1$ .

of the protein population, we provided explicit analytical conditions under which an auto-regulatory mechanism can evolve. Our conclusion is that negative feedback can evolve if and only if the quantity in equation (5.15) is positive.

A systematic analysis of (5.15) showed that a negative feedback is more likely to evolve when the source of fluctuations in the protein population is more due to extrinsic noise (i.e., from a noisy exogenous signal) than from intrinsic noise (i.e., from random protein expression and degradation events). This is exemplified by the fact that negative feedback with Hill coefficients in the range  $M_{ext} < M < M_{int}$  cannot evolve when protein noise is intrinsic but can evolve if there is sufficiently large extrinsic noise in protein numbers. The range of Hill coefficients which can only evolve in the presence of extrinsic noise, shrink as we decrease the response time of the exogenous signal (i.e., increase  $\kappa$ ). Combining the two previous observations we conclude that a



negative feedback is more likely to evolve when noise in protein numbers comes from a noisy exogenous signal with a large response time.

Figures 5.2 and 5.3 also show that evolvability of the negative feedback critically depends upon the threshold  $\mathbf{T}$ . Level of this critical threshold closer to the average population decrease the evolvability of the negative feedback. In particular, when the protein noise is dominated by extrinsic noise then we have from (5.17) that a feedback with  $M = 1$  can only evolve if the normalized threshold is less than  $1/2$  (when  $\kappa \ll 1$ ) and less than  $1/3$  (when  $\kappa \gg 1$ ).

Finally, our analysis also predicts that the evolvability of the negative feedback is independent of the constant  $b$  in the transcriptional response (5.8). This is simply because although a transcriptional response with a lower value of  $b$  causes a larger decrease in  $\mathbf{x}^*$  when a feedback is introduced, it is compensated by the fact that a lower  $b$  also causes a larger decrease in fluctuations about  $\mathbf{x}^*$ .

### 5.3.1 Future work

We analyzed the evolvability of the negative feedback using a simple form for the fitness function that is given by the probability of maintaining protein numbers above a critical level. A direction of future research is to modify this fitness function by assuming that there also exists an upper threshold  $\bar{\mathbf{T}}$ , and increasing protein numbers above  $\bar{\mathbf{T}}$  also compromises the cell's viability. Using this modified fitness function given by  $\text{Probability}\{\mathbf{T} < \mathbf{x} < \bar{\mathbf{T}}\}$  our goal would be to analyze how the evolvability of the feedback depends on  $\bar{\mathbf{T}}$ .

Recent work has introduced complex fitness functions for proteins that are involved in metabolizing sugars [2]. Another direction of future work is to extend the above

evolvability analysis to these fitness functions.

## Chapter 6

# Scaling of stochasticity in gene cascades

Gene cascades, where a protein expressed from one gene activates another gene to make a different protein, are common motifs occurring within cells. In these cascades an initial signal, which involves a protein with small number of molecules, can be amplified over a number of stages. The amplified signal can then be used to trigger some physiological response in the cell. We investigate how noise levels in the proteins of the gene cascade are effected by the number of stages and the per-stage magnification of the cascade.

We consider a cascade of  $Q$  genes  $GeneX_1, GeneX_2, \dots, GeneX_Q$  where gene  $GeneX_i$  expresses the protein  $X_i$ , which in turn activates gene  $GeneX_{i+1}$  to make the protein  $X_{i+1}$  for each  $i \in [1, \dots, Q-1]$ , as shown in Figure 6.1. In Section 6.1, we model the activation by assuming that the transcription rate of  $GeneX_{i+1}$  is given by  $r_{i+1}\mathbf{x}_i$  for some constant  $r_{i+1}$  where  $\mathbf{x}_i$  denotes the number of molecules of protein  $X_i$ .

We show in Section 6.2 that the steady-state noise in the protein  $X_i$ , which we denote by  $CV_{X_i}$ , decreases as the number of stages  $i$  increases and therefore the cascade acts like a *noise attenuator*, where downstream proteins have reduced noise.

We also compare the final protein noise in two different cascades of genes with different number of stages but with the same average steady-state number of molecules of the final protein. We show that the gene cascade that achieves the same total magnification with a large number of stages but smaller per-stage magnification exhibits lesser stochastic fluctuations in the final protein than the cascade with a smaller number of stages but a larger per-stage magnification.

In Section 6.3 we investigate the same cascade of  $Q$  genes, but now consider the situation where the transcription rate of  $GeneX_{i+1}$  is given by  $\mathbf{r}_{i+1}\mathbf{x}_i$  where now  $\mathbf{r}_{i+1}$  is a stochastic process. Fluctuations in  $\mathbf{r}_{i+1}$  are referred to as the extrinsic noise that enters the gene cascade. We show in Section 6.3 that the presence of this extrinsic noise can change the qualitative behaviour of the cascade. In particular, for large extrinsic noise,  $CV_{X_i}$  increases with the stage index  $i$ , and hence, the cascade behaves as a *noise magnifier* instead of a noise attenuator. We also show that, given two cascades with a different number of stages but the same average steady-state number of molecules of the final protein, there exists a critical level of extrinsic noise above which the cascade with the smaller number of stages and higher per-stage magnification exhibits lesser noise in the final protein than the cascade with a larger number of stages and lower per-stage magnification.

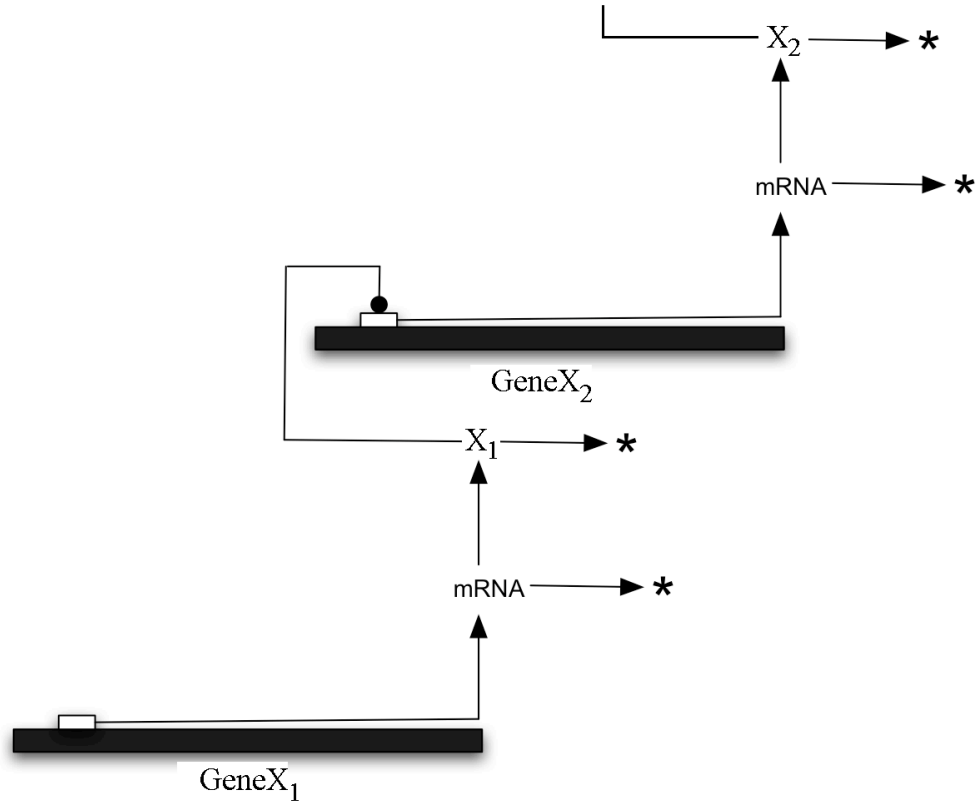


Figure 6.1. A gene activation cascade where gene  $GeneX_1$  expresses protein  $X_1$ . This protein then activates gene  $GeneX_2$  to make  $X_2$  which then goes on to activate gene  $GeneX_3$ .

## 6.1 Noise in a cascade of genes

We consider a cascade of  $Q$  genes  $GeneX_1, GeneX_2, \dots, GeneX_Q$  where gene  $GeneX_i$  expresses protein  $X_i$  that activates gene  $GeneX_{i+1}$ , for each  $i \in [1, \dots, Q-1]$  (as shown in Figure 6.1). We model the activation of  $GeneX_{i+1}$  by assuming that the transcription rate of  $GeneX_{i+1}$  is given by  $r_{i+1}\mathbf{x}_i$  for some constant  $r_{i+1}$  where  $\mathbf{x}_i$  denotes the number of molecules of protein  $X_i$ . The constant  $r_i$  typically depends on the gene copy number and on the number of molecules of transcription factors and enzymes involved in the expression of the corresponding gene. For now we take all the

$r_i$  to be constants but later we shall consider fluctuations in  $r_i$ .

As in Section 4.1, we consider a simple model of gene expression where expression of gene  $GeneX_i$  for all  $i \in [1, \dots, Q]$  leads to the formation of  $\mathbf{N}_i$  molecules of the protein  $X_i$  where  $\mathbf{N}_i$  is a random variable with mean  $N_i$  and variance  $V_i^2$ . We also assume that the protein  $X_i$  decays at a constant rate  $d_i$ . Using the results presented in Section 4.3 we can relate the average number of molecules and noise in protein  $X_{i+1}$  with that of protein  $X_i$ : For this cascade with a transcriptional response of the form  $r_{i+1}\mathbf{x}_i$ , this analysis leads to

$$\mathbf{x}_{i+1}^* = \frac{r_{i+1}N_{i+1}}{d_{i+1}}\mathbf{x}_i^* =: G_{i+1}\mathbf{x}_i^*, \quad \forall i \in [1, \dots, Q-1] \quad (6.1)$$

where  $\mathbf{x}_i^*$  denotes the steady-state average number of molecules of protein  $X_i$ . The constant  $G_{i+1} := r_{i+1}N_{i+1}/d_{i+1} > 1$  can be viewed as the multiplicative gain by which the population of protein  $X_i$  is magnified at the  $i+1^{th}$  stage of the cascade and is assumed to be larger than one.

The noise in protein  $X_{i+1}$  arises from two sources: fluctuations caused by random gene expression/protein degradation and fluctuations due to noise in the activating signal  $\mathbf{x}_i$ . Using (4.33) with the appropriate transcriptional response  $r_{i+1}\mathbf{x}_i$ , we conclude that

$$S = \frac{dr_{i+1}\mathbf{x}_i}{d\mathbf{x}_i} \frac{\mathbf{x}_i}{r_{i+1}\mathbf{x}_i} = 1 \quad (6.2)$$

which allow us to express the noise in protein  $X_{i+1}$  as a function of the noise in protein  $X_i$  according to

$$CV_{X_{i+1}}^2 = CV_{intX_{i+1}}^2 + \frac{d_{i+1}}{d_i + d_{i+1}} CV_{X_i}^2, \quad \forall i \in [1, \dots, Q-1] \quad (6.3)$$

where  $CV_{X_i}$  is the steady-state noise in the protein  $X_i$  and

$$CV_{intX_i}^2 = \frac{N_i^2 + V_i^2 + N_i}{2N_i\mathbf{x}_i^*}, \quad \forall i \in [1, \dots, Q] \quad (6.4)$$

represents the intrinsic noise in the protein  $X_i$ . For simplicity, we focus our attention on the biologically relevant case [24] where the number of proteins produced per mRNA follows a geometric distribution, in which case the variance  $V_i^2$  is equal to  $N_i^2 - N_i$  and

$$CV_{intX_i}^2 = \frac{N_i}{\mathbf{x}_i^*}, \quad \forall i \in [1, \dots, Q]. \quad (6.5)$$

## 6.2 Scaling of noise in gene cascades

In this section we investigate how the noise in protein numbers varies with the per-stage magnification and the number of stages of the cascade. Substituting (6.5) in (6.3) we conclude that

$$CV_{X_{i+1}}^2 = \frac{N_{i+1}}{\mathbf{x}_{i+1}^*} + \frac{d_{i+1}}{d_i + d_{i+1}} CV_{X_i}^2. \quad (6.6)$$

Assuming that for all  $i$ ,  $N_i$  is bounded from above and  $d_{i+1}/(d_i + d_{i+1})$  is strictly less than one, we have

$$\lim_{i \rightarrow \infty} CV_{X_i}^2 = 0. \quad (6.7)$$

Thus in this case the cascade acts like a noise attenuator, meaning that as we go down the cascade, the noise in the downstream proteins becomes smaller, as shown in Figure 6.2.

For a homogeneous cascade of genes where

$$N_i = N, \quad d_i = d, \quad r_i = r, \quad G_i = G = \frac{rN}{d} > 1, \quad (6.8)$$

and the transcription rate of the first gene is given by  $\lambda r$  for some constant  $\lambda$ , we conclude from (6.1) that

$$\mathbf{x}_i^* = \lambda G^i \quad (6.9)$$

and from (6.6) that

$$CV_{X_{i+1}}^2 = \frac{N}{\lambda G^{i+1}} + \frac{1}{2} CV_{X_i}^2 = \frac{N}{G\lambda} \frac{1}{2^i} \sum_{j=0}^i \left(\frac{2}{G}\right)^j = \frac{N((G/2)^{i+1} - 1)}{(G/2 - 1)G^{i+1}\lambda}. \quad (6.10)$$

Analysis of this expression shows that if the per stage magnification ( $G$ ) is greater than two, then  $CV_{X_i}^2$  decreases monotonically with  $i$ , as in Figure 6.2(a), while for  $G < 2$ ,  $CV_{X_i}^2$  first increases and then decreases, as in Figure 6.2(b).

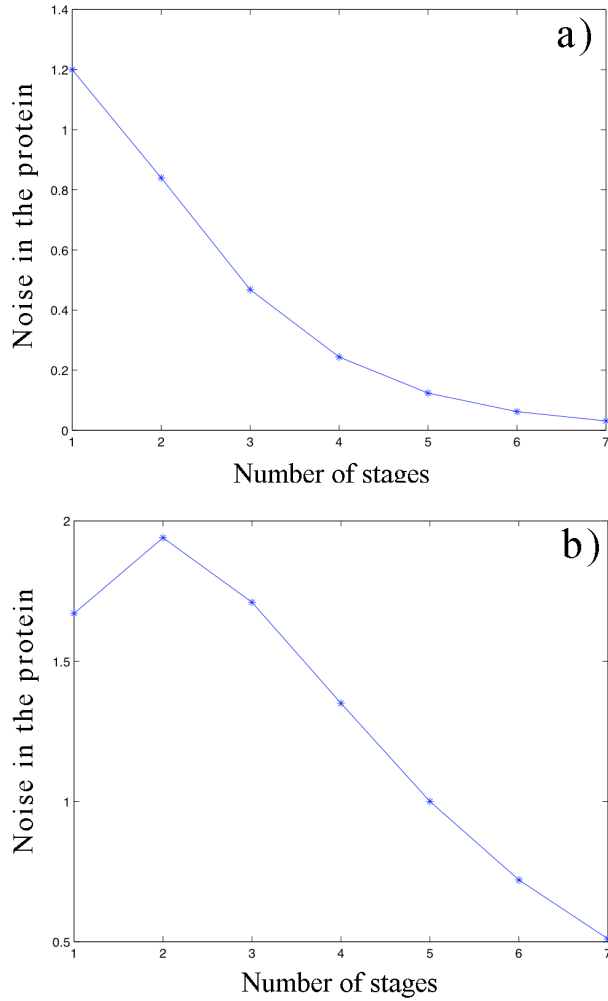


Figure 6.2.  $CV_{X_i}^2$  at the  $i^{th}$  stage of the gene cascade when all genes have per stage magnification of a)  $G = 5$  and b)  $G = 1.5$ . Other parameters taken as  $r = d = \lambda = 1$ .

We now compare the final protein noise in two different gene cascades with a



different number of stages but the same final average number of molecules. We assume that both the cascades have the same basic parameters, except for  $N$ . As  $G = rN/d$ , this implies that both cascades have different per stage magnifications  $G$ . For both cascades to have the same average number of molecules of the final protein, the cascade with a smaller number of stages will need to have a larger per-stage magnification. More specifically, if the cascades have  $V$  and  $W$  stages with per-stage magnification  $G_V$  and  $G_W$ , respectively, then

$$(G_V)^V = (G_W)^W. \quad (6.11)$$

From (6.10) we conclude that the ratio of the noises in the final protein are given by

$$\frac{CV_{X_V}^2}{CV_{X_W}^2} = \frac{G_V}{G_W} \frac{(G_W/2 - 1)}{(G_V/2 - 1)} \frac{((G_V/2)^V - 1)}{((G_W/2)^W - 1)} \quad (6.12)$$

where  $CV_{X_V}$  and  $CV_{X_W}$  denote the noise in the final protein of the gene cascades with  $V$  and  $W$  stages, respectively. If the per-stage magnification in both cascades is less than two then for large  $V$  and  $W$  we have

$$\frac{CV_{X_V}^2}{CV_{X_W}^2} \approx \frac{G_V}{G_W} \frac{(1 - G_W/2)}{(1 - G_V/2)}. \quad (6.13)$$

If  $V > W$ , which implies that  $G_V < G_W$ , then one can show that the above ratio is always smaller than one, i.e., the gene cascade with the larger number of stages exhibits lesser noise in the final protein. Alternatively, when the magnification is much larger than two, we conclude from (6.12) that

$$\frac{CV_{X_V}^2}{CV_{X_W}^2} \approx \frac{G_V}{G_W} \frac{G_W/2}{G_V/2} \frac{(G_V/2)^N}{(G_W/2)^K} = \left(\frac{1}{2}\right)^{V-W}. \quad (6.14)$$

This again shows that a gene cascade which achieves magnification with a larger number of stages but smaller per-stage magnification will exhibit lower stochastic fluctuations in the final protein than a cascade with a smaller number of stages but a larger per-stage magnification.

### 6.3 Scaling of extrinsic and intrinsic noise in gene cascades

In the previous section we assumed that the transcription rate of the gene  $GeneX_{i+1}$  is equal to  $r_{i+1}\mathbf{x}_i$  where  $r_{i+1}$  was assumed to be a fixed constant. We now consider the situation where  $r_{i+1}$  is not fixed but fluctuates about its mean value. More specifically, we assume the transcription rate of gene  $GeneX_{i+1}$  is  $\mathbf{r}_{i+1}\mathbf{x}_i$  where  $\mathbf{r}_{i+1}$  is a stochastic process with steady-state mean and coefficient of variation given by  $r_{i+1}$  and  $CV_{ext_{i+1}}$ , respectively. We refer to fluctuations in  $\mathbf{r}_i$  as the amount of extrinsic noise that enters the  $i^{th}$  stage of the cascade and denote it by  $CV_{ext_i}$ . For simplicity, we assume that  $\mathbf{r}_i$  and  $\mathbf{r}_j$  are independent of each other for  $i \neq j$ . Analysis in Appendix C shows that (6.1) still holds but now

$$CV_{X_{i+1}}^2 = \frac{N_{i+1}}{\mathbf{x}_{i+1}^*} + (\xi_i + \delta_i CV_{ext_i}^2) CV_{X_i}^2 + \zeta_i CV_{ext_i}^2 \quad (6.15)$$

where

$$\xi_i = \frac{d_{i+1}}{d_i + d_{i+1}} < 1, \quad \zeta_i = \frac{d_{i+1}}{g_{i+1} + d_{i+1}} < 1, \quad \delta_i = \frac{d_{i+1}}{g_{i+1} + d_{i+1} + d_i} < 1 \quad (6.16)$$

and  $g_i$  represents the degradation rate of the signal  $\mathbf{r}_{i+1}$ . As expected, when  $CV_{ext_i} = 0$  (6.15) reduces to (6.6).

One can now conclude from (6.15) that unlike in Section 6.2 where  $CV_{X_i}^2$  goes to zero for sufficiently large  $i$ ,  $CV_{X_i}^2$  now approaches some non-zero value determined by the amount of extrinsic noise in the cascade. Furthermore, when  $\xi_i + \delta_i CV_{ext_i}^2 > 1$ , for every  $i$ , then the gene cascade is a noise magnifier where the downstream proteins have increased noise. Hence, by altering the amount of extrinsic noise in the cascade one could change a cascade from being a noise attenuator (as in Section 6.2 where  $CV_{ext_i} = 0$ ) to a noise amplifier.

For a homogeneous cascade where

$$d_i = d, \quad r_i = r, \quad CV_{ext_i}^2 = CV_{ext}^2, \quad g_i = g, \quad N_i = N, \quad G_i = G = \frac{rN}{d} > 1, \quad (6.17)$$

with the transcription rate of the first gene given by  $\lambda r$  for some constant  $\lambda$ , we have from (6.15) that

$$CV_{X_{i+1}}^2 = \frac{N}{\lambda G^{i+1}} + \frac{d}{d+g} CV_{ext}^2 + \left( \frac{1}{2} + \frac{d}{2d+g} CV_{ext}^2 \right) CV_{X_i}^2. \quad (6.18)$$

We conclude from (6.18) that

$$\lim_{i \rightarrow \infty} CV_{X_i}^2 = \begin{cases} S & \text{for } CV_{ext}^2 < 1 + \frac{g}{2d} \\ \infty & \text{for } CV_{ext}^2 \geq 1 + \frac{g}{2d} \end{cases} \quad (6.19)$$

where

$$S = \frac{d CV_{ext}^2}{d+g} \frac{2(2d+g)}{2d(1-CV_{ext}^2)+g}. \quad (6.20)$$

We can explicitly solve the difference equation (6.18) to obtain the noise in protein  $X_i$  to be

$$CV_{X_i}^2 = \frac{dG((G\alpha)^i - 1)}{r(G\alpha - 1)\lambda G^i} + \zeta CV_{ext}^2 \frac{\alpha^i - 1}{\alpha - 1} \quad (6.21)$$

where

$$\alpha = \frac{1}{2} + \frac{d}{2d+g} CV_{ext}^2, \quad \zeta = \frac{d}{d+g}. \quad (6.22)$$

For sufficiently large  $G$  the above expression reduces to

$$CV_{X_i}^2 \approx \frac{d\alpha^{i-1}}{r\lambda} + \zeta CV_{ext}^2 \frac{\alpha^i - 1}{\alpha - 1}. \quad (6.23)$$

As in the previous section we now consider two cascades of lengths  $V, W$  with  $V > W$  and per-stage magnifications  $G_V$  and  $G_W$  chosen such that both cascades result in the

same average number of molecules for the final protein (i.e., (6.11) holds). We showed in Section 6.2 that if  $CV_{ext} = 0$  then  $CV_{X_V}^2 < CV_{X_W}^2$ . However, from (6.23) we conclude that when  $\alpha = 1$ , i.e.,  $CV_{ext}^2 = 1 + g/2d$  we have

$$CV_{X_i}^2 \approx \frac{d}{r\lambda} + i\zeta \left(1 + \frac{g}{2d}\right). \quad (6.24)$$

Thus when  $CV_{ext}^2 = 1 + g/2d$ ,  $CV_{X_V}^2 > CV_{X_W}^2$  as  $V > W$ . Since  $CV_{X_i}^2$  varies continuously with  $CV_{ext}^2$ , there must exist a critical value  $CV_{crit} < \sqrt{1 + g/2d}$  for the extrinsic noise  $CV_{ext}$  such that when  $CV_{ext} > CV_{crit}$ , then  $CV_{X_V} > CV_{X_W}$ . This shows that when the extrinsic noise is larger than a critical value, then the gene cascade which achieves the same magnification with a smaller number of stages and higher per-stage magnification exhibits lower stochastic fluctuations in the final protein than the cascade with a larger number of stages and a smaller per-stage magnification.

## 6.4 Conclusion

This chapter presented results relating the stochastic noise for proteins in a gene cascade with the number of stages and the per-stage magnification. We provided explicit formulas to compute the noise in the proteins both in the absence and presence of extrinsic noise. We showed that when there is no extrinsic noise the noise in the  $i^{th}$  protein decreases with  $i$  and the cascade acts like a *noise attenuator*. Furthermore, for two different cascades with the same average final protein level, the cascade with the larger number of stages exhibits lower stochastic noise in the final protein. However, in the presence of a sufficiently large level of extrinsic noise there is a role reversal for the cascade. More specifically, the cascade now acts as a *noise amplifier* where downstream proteins have increased noise. Thus depending on whether noise in the final

product is deleterious or advantageous to the cell, it can be appropriately modulated via cascades.

### **6.4.1 Future work**

We assumed in Section 6.3 that  $\mathbf{r}_i$  and  $\mathbf{r}_j$  for  $i \neq j$  are independent of each other. This would be a valid approximation if plasmids encoding different genes fluctuate independently of each other. However, if the source of extrinsic noise are fluctuations in enzyme levels then  $\mathbf{r}_i$  and  $\mathbf{r}_j$  for  $i \neq j$  would be positively correlated as generally the same enzymes are involved in the transcription of the gene. A direction of future work is to investigate scenarios where correlations exists between  $\mathbf{r}_i$  and study its consequences on the scaling of stochasticity in cascades.

# Bibliography

- [1] Murat Acar, Attila Becskei, and Alexander van Oudenaarden. Enhancement of cellular memory by reducing stochastic transitions. *Nature*, 435, 2005.
- [2] U. Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC, 2006.
- [3] U. Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8:450–461, 2007.
- [4] A.P. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage  $\lambda$ -infected *Escherichia coli* cells. *Genetics*, 149:1633–1648, 1998.
- [5] N. T. J. Bailey. *The Elements of Stochastic Processes*. Wiley, New York, 1964.
- [6] Attila Becskei and Luis Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405:590–593, 2000.
- [7] Martin Bengtsson, Anders Stahlberg, Patrik Rorsman, and Mikael Kubista. Gene expression profiling in single cells from the pancreatic islets of langerhans reveals lognormal distribution of mRNA levels. *Genome Res.*, 15:1388–1392, 2005.

- [8] William J. Blake, Mads Kaern, Charles R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422:633–637, 2003.
- [9] Onn Brandman, James E. Ferrell, Rong Li, and Tobias Meyer. Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science*, 310:496 – 498, 2005.
- [10] M. H. A. Davis. *Markov models and Optimization*. Chapman and Hall, 1993.
- [11] Yann Dublanche, Konstantinos Michalodimitrakis, Nico Kummerer, Mathilde Foglierini, and Luis Serrano. Noise in transcription negative feedback loops: simulation and experimental analysis. *Molecular Systems Biology*, 2006.
- [12] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. Stochastic gene expression in a single cell. *Science*, 297:1183–1186, 2002.
- [13] Hunter B. Fraser, Aaron E. Hirsh, Guri Giaever, Jochen Kumm, and Michael B. Eisen. Noise minimization in eukaryotic gene expression. *PLoS Biology*, 2, 2004.
- [14] Daniel T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. of Computational Physics*, 22:403–434, 1976.
- [15] Daniel T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. of Chemical Physics*, 115(4):1716–1733, 2001.
- [16] C. A. Gomez-Urbe and G. C. Verghese. Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *J. of Chemical Physics*, 126, 2007.

- [17] J. Hasty, J. Pradines, M. Dolnik, and J. J. Collins. Noise-based switches and amplifiers for gene expression. *Proceedings of the National Academy of Sciences*, 97:2075–2080, 2000.
- [18] João Pedro Hespanha. Stochastic hybrid systems: Applications to communication networks. In Rajeev Alur and George J. Pappas, editors, *Hybrid Systems: Computation and Control*, number 2993 in Lect. Notes in Comput. Science, pages 387–401. Springer-Verlag, Berlin, March 2004.
- [19] João Pedro Hespanha and Abhyudai Singh. Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems. *Int. J. of Robust and Nonlinear Control*, 15:669–689, 2005.
- [20] Sara Hooshangi and Ron Weiss. The effect of negative feedback on noise propagation in transcriptional gene networks. *CHAOS*, 16, 2006.
- [21] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier Science, Amsterdam, The Netherlands, 2001.
- [22] M. J. Keeling. Multiplicative moments and measures of persistence in ecology. *J. of Theoretical Biology*, 205:269–281, 2000.
- [23] Isthinayagi Krishnarajah, Alex Cook, Glenn Marion, and Gavin Gibson. Novel moment closure approximations in stochastic epidemics. *Bulletin of Mathematical Biology*, 67:855–873, 2005.
- [24] Diane Longo and Jeff Hasty. Imaging gene expression: Tiny signals make a big noise. *Nature Chemical Biology*, 2:181–182, 2006.
- [25] J. H. Matis and T. R. Kiffe. On approximating the moments of the equilibrium distribution of a stochastic logistic model. *Biometrics*, 52:155–166, 1996.



- [26] J. H. Matis and T. R. Kiffe. On interacting bee/mite populations: a stochastic model with analysis using cumulant truncation. *Enviromental and Ecological Statistics*, 9:237–258, 2002.
- [27] J. H. Matis, T. R. Kiffe, and P. R. Parthasarathy. On the cumulant of population size for the stochastic power law logisitc model. *Theoretical Population Biology*, 53:16–29, 1998.
- [28] H. H. McAdams and A. P. Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94:814–819, 1997.
- [29] D. A. McQuarrie. Stochastic approach to chemical kinetics. *J. of Applied Probability*, 4:413–478, 1967.
- [30] Brian Munsky and Mustafa Khammash. The finite state projection algorithm for the solution of the chemical master equation. *J. of Chemical Physics*, 124, 2006.
- [31] I. Nasell. Extinction and quasi-stationarity in the verhulst logistic model. *J. of Theoretical Biology*, 211:11–27, 2001.
- [32] I. Nasell. An extension of the moment closure method. *Theoretical Population Biology*, 64:233–239, 2003.
- [33] I. Nasell. Moment closure and the stochastic logistic model. *Theoretical Population Biology*, 63:159–168, 2003.
- [34] A. S. Novozhilov, G. P. Karev, and E. V. Koonim. Biological applications of the theory of birth-and-death processes. *Briefings in Bioinformatics*, 7(1):70–85, 2006.

- [35] David Orrell and Hamid Bolouri. Control of internal and external noise in genetic regulatory networks. *J. of Theoretical Biology*, 230:301–312, 2004.
- [36] Ertugrul M. Ozbudak, Mukund Thattai, Iren Kurtser, Alan D. Grossman, and Alexander van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature Genetics*, 31:69–73, 2002.
- [37] Johan Paulsson. Summing up the noise in gene networks. *Nature*, 427:415–418, 2004.
- [38] Johan Paulsson. Model of stochastic gene expression. *Physics of Life Reviews*, 2:157–175, 2005.
- [39] Johan Paulsson, Otto G. Berg, and Mans Ehrenberg. Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proceedings of the National Academy of Sciences*, 97(13):7148–7153, 2000.
- [40] R. Pearl and L. J. Reed. On the rate of growth of the population of the united states since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences*, 6:275–288, 1920.
- [41] E. C. Pielou. *Mathematical Ecology*. Wiley, New York, 1977.
- [42] Mark Ptashne. *Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2004.
- [43] C. V. Rao and A. P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the gillespie algorithm. *J. of Chemical Physics*, 118:4999–5010, 2003.

- [44] Jonathan M. Raser and Erin K. O’Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309:2010 – 2013, 2005.
- [45] N. Rosenfeld, M. B. Elowitz, and U. Alon. Negative autoregulation speeds the response times of transcription networks. *J. Molecular Biology*, 323:785–793, 2002.
- [46] Michael A. Savageau. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature*, 252:546–549, 1974.
- [47] Vahid Shahrezaei, Julien F. Ollivier, and Peter S. Swain. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology*, 4, 2008.
- [48] Michael L. Simpson, Chris D. Cox, and Gary S. Sayler. Frequency domain analysis of noise in autoregulated gene circuits. *PNAS*, 100:4551–4556, 2003.
- [49] Abhyudai Singh and João Pedro Hespanha. Lognormal moment closures for biochemical reactions. In *Proc. of the 45th Conf. on Decision and Control, San Diego*, 2006.
- [50] Abhyudai Singh and João Pedro Hespanha. Moment closure techniques for stochastic models in population biology. In *Proc. of the 2006 Amer. Control Conference, Minneapolis, MN*, 2006.
- [51] Abhyudai Singh and João Pedro Hespanha. Stochastic analysis of gene regulatory networks using moment closure. In *Proc. of the 2007 Amer. Control Conference, New York, NY*, 2006.
- [52] Kim Sneppen, Mille A. Micheelsen, and Ian B. Dodd. Ultrasensitive gene regulation by positive feedback loops in nucleosome modification. *Molecular Systems Biology*, 4(182), 2008.

- [53] John L. Spudich and D. E. Koshland Jr. Non-genetic individuality: chance in the single cell. *Nature*, 262:467–471, 1976.
- [54] R. Srivastava, L. You, J. Summers, and J. Yin. Stochastic vs. deterministic modeling of intracellular viral kinetics. *J. of Theoretical Biology*, 218(3):309–321, 2002.
- [55] Dov J. Stekel and Dafyd J. Jenkins. Strong negative self regulation of prokaryotic transcription factors increases the intrinsic noise of protein expression. *BMC Systems Biology*, 2008.
- [56] Peter S. Swain, Michael B. Elowitz, and Eric D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *PNAS*, 99:12795–12800, 2002.
- [57] Yi Tao, Xiudeng Zheng, and Yuehua Sun. Effect of feedback regulation on stochastic gene expression. *J. of Theoretical Biology*, 247:827–836, 2007.
- [58] Mukund Thattai and Alexander van Oudenaarden. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15):8614–8619, 2001.
- [59] Ryota Tomioka, Hidenori Kimura, Tetsuya J. Kobayashi, and Kazuyuki Aihara. Multivariate analysis of noise in genetic regulatory networks. *J. of Theoretical Biology*, 229:501–521, 2004.
- [60] Michael E. Wall, William S. Hlavacek, and Michael A. Savageau. Design principles for regulator gene expression in a repressible gene circuit. *J. of Molecular Biology*, 332:861–876, 2003.
- [61] M. C. Walters, S. Fiering, J. Eidemiller, W. Magis, M. Groudine, and D. I. K.

- Martin. Enhancers increase the probability but not the level of gene expression. *Proceedings of the National Academy of Sciences*, 92:7125–7129, 1995.
- [62] L. Weinberger, Roy D. Dar, and Michael L. Simpson. Transient-mediated fate determination in a transcriptional circuit of hiv. *Nature Genetics*, 40:466 – 470, 2008.
- [63] L. Weinberger and Thomas Shenk. An hiv feedback resistor: Auto-regulatory circuit deactivator and noise buffer. *PLoS Biology*, 5, 2007.
- [64] P. Whittle. On the use of the normal approximation in the treatment of stochastic processes. *J. Roy. Statist. Soc., Ser. B*, 19:268–281, 1957.
- [65] Frank Wilkinson. *Chemical Kinetics and Reaction Mechanisms*. Van Nostrand Reinhold Co, New York, 1980.

# Appendix A

## Moment closure for chemical reactions

To evaluate the time evolution of moments we use the fact (see [10, 18] for details) that for the SHS (1.6)-(1.8), the time evolution of any differentiable function  $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$  is given by

$$\frac{d\mathbf{E}[\psi(\mathbf{x})]}{dt} = \mathbf{E}[(\mathbf{L}\psi)(\mathbf{x})], \quad (\text{A.1})$$

where  $\forall \mathbf{x} \in \mathbb{R}^n$

$$(\mathbf{L}\psi)(\mathbf{x}) := \sum_{i=1}^K (\psi(\phi_i(\mathbf{x})) - \psi(\mathbf{x})) \lambda_i(\mathbf{x}). \quad (\text{A.2})$$

Let  $\mu_{(\mathbf{m})}$  be a  $m^{th}$  order moment of  $\mathbf{x}$ . Using (1.9) and replacing  $\psi(\mathbf{x})$  with  $\mathbf{x}^{(\mathbf{m})}$  in (A.1) and (A.2), we have that

$$\frac{d\mu_{(\mathbf{m})}}{dt} = \mathbf{E}[(\mathbf{L}\psi)(\mathbf{x})] \quad (\text{A.3})$$

where

$$(\mathbf{L}\psi)(\mathbf{x}) = \sum_{i=1}^K c_i h_i(\mathbf{x}) \left\{ \phi_i(\mathbf{x})^{(\mathbf{m})} - \mathbf{x}^{(\mathbf{m})} \right\} \quad (\text{A.4a})$$

$$= \sum_{i=1}^K c_i h_i(\mathbf{x}) \left\{ \left[ \prod_{j=1}^n (\mathbf{x}_j - u_{ij} + v_{ij})^{m_j} \right] - \mathbf{x}^{(\mathbf{m})} \right\} \quad (\text{A.4b})$$

$$= \sum_{i=1}^K c_i h_i(\mathbf{x}) \left\{ \left[ \prod_{j=1}^n \sum_{q=0}^{m_j} \mathbf{C}_q^{m_j} \mathbf{x}_j^{m_j-q} a_{ij}^q \right] - \mathbf{x}^{(\mathbf{m})} \right\} \quad (\text{A.4c})$$

and  $a_{ij} := v_{ij} - \mu_{ij}$ . As  $(\mathbf{L}\psi)(\mathbf{x})$  is simply a linear combination of monomials of  $\mathbf{x}$ , we have from (A.3) that the time derivative of  $\mu_{(\mathbf{m})}$  is a linear combination of moments.

Moreover as

$$(\mathbf{L}\psi)(\mathbf{x}) = \sum_{i=1}^K c_i h_i(\mathbf{x}) \left\{ \left[ \prod_{j=1}^n \sum_{q=0}^{m_j} \mathbf{C}_q^{m_j} \mathbf{x}_j^{m_j-q} a_{ij}^q \right] - \mathbf{x}^{(\mathbf{m})} \right\} \quad (\text{A.5a})$$

$$= \sum_{i=1}^K c_i h_i(\mathbf{x}) \left\{ \sum_{j=1}^n \mathbf{C}_1^{m_j} \frac{\mathbf{x}^{(\mathbf{m})}}{\mathbf{x}_j} a_{ij} + \dots \right\}, \quad (\text{A.5b})$$

the highest order moment appearing in  $\dot{\mu}_{(\mathbf{m})}$  is determined by the degree of the polynomial

$$\sum_{i=1}^K c_i h_i(\mathbf{x}) \sum_{j=1}^n \mathbf{C}_1^{m_j} a_{ij} \frac{\mathbf{x}^{(\mathbf{m})}}{\mathbf{x}_j}. \quad (\text{A.6})$$

As  $\mathbf{x}^{(\mathbf{m})}$  is a monomial of degree  $m$ , and  $h_i(\mathbf{x})$  can be a polynomial of degree at most 2, the polynomial (A.6) is a polynomial in  $\mathbf{x}$  of degree at most  $m+1$ . Thus the highest order moment that can appear in the time derivative of  $\mu_{(\mathbf{m})}$  is of order  $m+1$ . Hence, if one constructs a vector

$$\boldsymbol{\mu} = [\mu_{(\mathbf{m}_1)}, \dots, \mu_{(\mathbf{m}_k)}]^T \in \mathbb{R}^k, \quad \mathbf{m}_p \in \mathbb{N}_{\geq 0}^n, \quad \forall p \in \{1, \dots, k\} \quad (\text{A.7})$$

containing all the moments of  $\mathbf{x}$  of order upto  $\mathbf{R}$ , then its time derivative is given by

$$\dot{\boldsymbol{\mu}} = \hat{\mathbf{a}} + A\boldsymbol{\mu} + B\bar{\boldsymbol{\mu}} \quad (\text{A.8})$$

where  $\bar{\mu} \in \mathbb{R}^r$  is a vector of moments of order  $\mathbf{R} + 1$ . Let  $\bar{n} = k - n^* + 1$ , where  $n^*$  denotes the row in the vector  $\mu$  from where the  $\mathbf{R}^{th}$  order moments of  $\mathbf{x}$  start appearing. Then all elements in the first  $n^* - 1$  rows of matrix  $B$  will be zero. This is because the time derivative of a moment of order less than  $\mathbf{R}$ , is dependent at most on moment of order up to  $\mathbf{R}$  which are already present in the vector  $\mu$ .



# Appendix B

## Optimal feedback strength

### B.1 Effects of nonlinearities in $g(\mathbf{x})$

For the transcriptional response given by (4.24) the time evolution of the moments  $\mathbf{E}[\mathbf{x}]$ ,  $\mathbf{E}[\mathbf{x}^2]$  are given by

$$\begin{aligned} \frac{d\mathbf{E}[\mathbf{x}]}{dt} = & g(\mathbf{x}^*)N_x - d_x\mathbf{E}[\mathbf{x}] + g'(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}] + \frac{g''(\mathbf{x}^*)}{2}N_x\mathbf{E}[\mathbf{x}^2] - g'(\mathbf{x}^*)N_x\mathbf{x}^* \\ & - g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}]\mathbf{x}^* + \frac{g''(\mathbf{x}^*)}{2}N_x\mathbf{x}^{*2} \end{aligned} \quad (\text{B.1a})$$

$$\begin{aligned} \frac{d\mathbf{E}[\mathbf{x}^2]}{dt} = & g(\mathbf{x}^*)(N_x^2 + V_x^2) + d_x\mathbf{E}[\mathbf{x}] + 2g(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}] + g'(\mathbf{x}^*)(N_x^2 + V_x^2)\mathbf{E}[\mathbf{x}] \\ & - 2d_x\mathbf{E}[\mathbf{x}^2] + 2g'(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}^2] + \frac{g''(\mathbf{x}^*)}{2}(N_x^2 + V_x^2)\mathbf{E}[\mathbf{x}^2] \\ & + g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}^3] - g'(\mathbf{x}^*)(N_x^2 + V_x^2)\mathbf{x}^* - 2g'(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}]\mathbf{x}^* \\ & - g''(\mathbf{x}^*)(N_x^2 + V_x^2)\mathbf{E}[\mathbf{x}]\mathbf{x}^* - 2g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}^2]\mathbf{x}^* \\ & + \frac{g''(\mathbf{x}^*)}{2}(N_x^2 + V_x^2)\mathbf{x}^{*2} + g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}]\mathbf{x}^{*2} \end{aligned} \quad (\text{B.1b})$$

which can be written more compactly as

$$\begin{bmatrix} \frac{d\mathbf{E}[\mathbf{x}]}{dt} \\ \frac{d\mathbf{E}[\mathbf{x}^2]}{dt} \end{bmatrix} = \mathbf{a} + \mathbf{A} \begin{bmatrix} \mathbf{E}[\mathbf{x}] \\ \mathbf{E}[\mathbf{x}^2] \end{bmatrix} + \mathbf{B} \mathbf{E}[\mathbf{x}^3], \quad (\text{B.2})$$

for some vector  $\mathbf{a}$  and matrices  $\mathbf{A}$  and  $\mathbf{B}$ . The above moment dynamics is not closed as the time derivative of  $\mathbf{E}[\mathbf{x}]$  and  $\mathbf{E}[\mathbf{x}^2]$  depends on  $\mathbf{E}[\mathbf{x}]$ ,  $\mathbf{E}[\mathbf{x}^2]$  and  $\mathbf{E}[\mathbf{x}^3]$ . We use moment closure techniques to approximate  $\mathbf{E}[\mathbf{x}^3]$  as a nonlinear function of  $\mathbf{E}[\mathbf{x}]$  and  $\mathbf{E}[\mathbf{x}^2]$  and close the above set of differential equations. A standard assumption at this point is to assume that the third cumulant of the distribution is zero, which will be valid approximation as long as the distribution is symmetrically distributed about its mean. Referring the reader to [16, 49, 33] for further details we approximate  $\mathbf{E}[\mathbf{x}^3]$  as

$$\mathbf{E}[\mathbf{x}^3] \approx 3\mathbf{E}[\mathbf{x}]\mathbf{E}[\mathbf{x}^2] - 2\mathbf{E}[\mathbf{x}]^3. \quad (\text{B.3})$$

Denoting the steady-state values of the moments  $\mathbf{E}[\mathbf{x}]$  and  $\mathbf{E}[\mathbf{x}^2]$  by  $\mathbf{x}_q^*$  and  $\mathbf{E}^*[\mathbf{x}^2]$ , respectively, we have from (B.2) and (B.3) that

$$0 = \mathbf{a} + \mathbf{A} \begin{bmatrix} \mathbf{x}_q^* \\ \mathbf{E}^*[\mathbf{x}^2] \end{bmatrix} + \mathbf{B} \left( 3\mathbf{x}_q^* \mathbf{E}^*[\mathbf{x}^2] - 2\mathbf{x}_q^{*3} \right). \quad (\text{B.4})$$

Analytically solving for these steady-state moments from (B.4) is not an easy task and we use perturbation methods to compute approximate steady-states. This is done by writing  $\mathbf{x}_q^*$  as a perturbation about  $\mathbf{x}^*$  and  $\mathbf{E}^*[\mathbf{x}^2]$  as a perturbation about  $\mathbf{x}_q^{*2}$ , as follows

$$\mathbf{x}_q^* := \mathbf{x}^*(1 + \varepsilon_1), \quad \mathbf{E}^*[\mathbf{x}^2] := \mathbf{x}_q^{*2}(1 + \varepsilon_2). \quad (\text{B.5})$$

Assuming

$$\varepsilon_1^2 \ll 1, \quad \varepsilon_2^2 \ll 1, \quad |\varepsilon_1 \varepsilon_2| \ll 1 \quad (\text{B.6})$$

we have from (B.5)

$$\mathbf{E}^*[\mathbf{x}^2] \approx \mathbf{x}^{*2}(1 + 2\varepsilon_1 + \varepsilon_2) \quad (\text{B.7a})$$

$$3\mathbf{x}_q^* \mathbf{E}^*[\mathbf{x}^2] - 2\mathbf{x}_q^{*3} \approx \mathbf{x}^{*3}(1 + 3\varepsilon_1 + 3\varepsilon_2). \quad (\text{B.7b})$$

Substituting (B.7) in (B.4) we obtain

$$0 = \tilde{\mathbf{a}} + \tilde{\mathbf{A}} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad (\text{B.8})$$

for some vector  $\tilde{\mathbf{a}}$  and matrix  $\tilde{\mathbf{A}}$ . Solving for  $\varepsilon_1$  and  $\varepsilon_2$  from (B.8) we have

$$\varepsilon_1 = \frac{1}{1 + \frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV_{int}^2}{2\lambda}} - 1, \quad (\text{B.9a})$$

$$CV_{int-quad}^2 = \frac{\mathbf{E}^*[\mathbf{x}^2] - \mathbf{x}^{*2}}{\mathbf{x}_q^{*2}} = \varepsilon_2 = \frac{CV_{int}^2}{1 + \frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV_{int}^2}{2\lambda}} \quad (\text{B.9b})$$

where  $CV_{int}^2$  is given by (4.20).

## B.2 Extrinsic and intrinsic contributions of noise

We model the time evolution of the number of molecules  $\mathbf{x}$  and  $\mathbf{z}$  through a Stochastic Hybrid System (SHS) with state  $\mathbf{y} = [\mathbf{z}, \mathbf{x}]^T$  characterized by trivial continuous dynamics

$$\dot{\mathbf{y}} = \begin{bmatrix} \dot{\mathbf{z}} \\ \dot{\mathbf{x}} \end{bmatrix} = 0 \quad (\text{B.10a})$$

and four reset maps

$$\mathbf{y} \mapsto \phi_1(\mathbf{y}) = \begin{bmatrix} \mathbf{z} + \mathbf{N}_z \\ \mathbf{x} \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_2(\mathbf{y}) = \begin{bmatrix} \mathbf{z} - 1 \\ \mathbf{x} \end{bmatrix} \quad (\text{B.10b})$$

$$\mathbf{y} \mapsto \phi_3(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x} + \mathbf{N}_x \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_4(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x} - 1 \end{bmatrix} \quad (\text{B.10c})$$

with corresponding transition intensities given by

$$\lambda_1(\mathbf{y}) = K_z, \quad \lambda_2(\mathbf{y}) = d_z \mathbf{z}, \quad (\text{B.11a})$$

$$\lambda_3(\mathbf{y}) = g(\mathbf{x}^*, \mathbf{z}^*) + \frac{dg(\mathbf{x}, \mathbf{z}^*)}{d\mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}^*} (\mathbf{x} - \mathbf{x}^*) + \frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}} \Big|_{\mathbf{z}=\mathbf{z}^*} (\mathbf{z} - \mathbf{z}^*), \quad (\text{B.11b})$$

$$\lambda_4(\mathbf{y}) = d_x \mathbf{x}. \quad (\text{B.11c})$$

Using (4.8) we have that the time evolution of all the first and second order moments of  $\mathbf{y}$  is given by

$$\begin{bmatrix} \frac{d\mathbf{E}[\mathbf{z}]}{dt} \\ \frac{d\mathbf{E}[\mathbf{x}]}{dt} \\ \frac{d\mathbf{E}[\mathbf{z}^2]}{dt} \\ \frac{d\mathbf{E}[\mathbf{x}^2]}{dt} \\ \frac{d\mathbf{E}[\mathbf{zx}]}{dt} \end{bmatrix} = \bar{a} + \bar{A} \begin{bmatrix} \mathbf{E}[\mathbf{z}] \\ \mathbf{E}[\mathbf{x}] \\ \mathbf{E}[\mathbf{z}^2] \\ \mathbf{E}[\mathbf{x}^2] \\ \mathbf{E}[\mathbf{zx}] \end{bmatrix} \quad (\text{B.12})$$

for some vector  $\bar{a}$  and matrix  $\bar{A}$ . A steady-state analysis of (B.12) shows that the steady-state average number molecules of the protein,  $\mathbf{x}^*$ , is given as the solution to (4.16) and the total noise level in the protein population is given by (4.33).

### B.3 Limit of noise suppression

It is not easy to derive an explicit expression for the minimum protein noise,  $CV_{tot-min}$ . However, for the biologically meaningful case of

$$b \ll 1, \quad T_z \gg T_{nr}, \quad (\text{B.13})$$

analytical formulations for both  $CV_{tot-min}$  and the optimal level of feedback strength which achieves this minimum noise are possible. When these assumptions are true we have from (4.53) that the total protein noise level is given by

$$CV_{tot}^2 = CV_{int-nr}^2 \frac{(1 + (a\mathbf{x}^*)^M)^2}{1 + (M+1)(a\mathbf{x}^*)^M} + S^2 CV_z^2 \left( \frac{1 + (a\mathbf{x}^*)^M}{1 + (M+1)(a\mathbf{x}^*)^M} \right)^2. \quad (\text{B.14})$$

Straightforward calculus shows that the right-hand-side of (B.14) is minimum when

$$(a\mathbf{x}^*)^M = \frac{M - 2 + \sqrt{M} \sqrt{8S^2 CV_z^2 + M CV_{int-nr}^2 / CV_{int-nr}}}{2(M+1)}, \quad (\text{B.15})$$

which implies from (B.14) and (4.38) that

$$\begin{aligned} CV_{tot-min}^2 &= \frac{2L}{(1+M)^2 \left( \sqrt{M} CV_{int-nr} + \sqrt{8S^2 CV_z^2 + M CV_{int-nr}^2} \right)^2} \\ L &= 4S^4 CV_z^4 + 19MS^2 CV_z^2 CV_{int-nr}^2 + 4M^2 CV_{int-nr}^4 \\ &\quad + 5\sqrt{M} S^2 CV_z^2 CV_{int-nr} \sqrt{8S^2 CV_z^2 + M CV_{int-nr}^2} \\ &\quad + 4\sqrt{M^3} CV_{int-nr}^3 \sqrt{8S^2 CV_z^2 + M CV_{int-nr}^2} \end{aligned} \quad (\text{B.16})$$

and

$$a_{min} = \frac{d_x \sqrt{M}}{2g_0 N_x (M+1)} \left( 3\sqrt{M} + \sqrt{8S^2 CV_z^2 + M CV_{int-nr}^2} / CV_{int-nr} \right) P \quad (\text{B.17a})$$

$$P = \left( \frac{M - 2 + \sqrt{M} \sqrt{8S^2 CV_z^2 + M CV_{int-nr}^2} / CV_{int-nr}}{2(M+1)} \right)^{\frac{1}{M}}. \quad (\text{B.17b})$$

In situations where the assumptions listed in (B.13) do not hold, both  $CV_{tot-min}$  and  $a_{min}$  can be obtained numerically by minimizing the right-hand-side of (4.53).

From (4.55) we conclude that when  $T_z \gg T_{nr}$ , the protein noise  $CV_{tot-nr}$  when there is no feedback is

$$CV_{tot-nr}^2 = CV_{int-nr}^2 + S^2 CV_z^2. \quad (\text{B.18})$$

Hence, given experimental measurements of  $CV_{tot-nr}$  and the minimal noise  $CV_{tot-min}$  in the protein population, one can determine  $CV_z$  by simultaneously solving (B.16) and (B.18). In cases where only  $CV_{tot-min}$  is obtained experimentally, then given an estimate of  $CV_{int-nr}$ , one can compute  $CV_z$  from (B.16).

In addition to the above assumptions (i.e.,  $b \ll 1$  and  $T_z \gg T_{nr}$ ) if we also have that

$$CV_{int-nr}^2 \ll S^2 CV_z^2, \quad (\text{B.19})$$

then (B.16) reduce to

$$CV_{tot-min}^2 \approx \frac{S^2 CV_z^2}{(1+M)^2} + \frac{5S CV_z CV_{int-nr} \sqrt{M}}{\sqrt{2}(1+M)^2}. \quad (\text{B.20})$$

## B.4 Estimating the noise in the exogenous signal

Assuming the source of extrinsic noise to be the plasmid population, we have from Section 4.3 that  $g(\mathbf{x}, \mathbf{z}) = \mathbf{z}g(\mathbf{x})$  and therefore

$$S = \frac{\mathbf{z}^*}{g(\mathbf{x}^*, \mathbf{z}^*)} \frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}} \Big|_{\mathbf{z}=\mathbf{z}^*} = 1. \quad (\text{B.21})$$

For this synthetic auto-regulatory gene network  $b = 0$  and we calculate from the reaction rates provided in Table I of [11],

$$CV_{int-nr}^2 \approx 0.008, \quad T_{nr}/T_z \approx 0.1. \quad (\text{B.22})$$

As the assumptions listed in (B.13) hold, we use the formulas in Appendix B.3 to quantify the noise  $CV_z$  in the exogenous signal. Using  $M = 1$ , the above estimate of  $CV_{int-nr}$ , and the experimentally obtained value of  $CV_{tot-min} \approx 0.4$ , we obtain from (B.20) that  $CV_z = CV_{plasmid}$  is approximately 0.64.

## B.5 Gene expression model with transcription and translation

In this section we consider a model of gene expression which takes into account the mRNA dynamics. Here the protein production is decomposed into two steps: transcription and translation (as shown in Figure B.1). We assume that the mRNA is tran-

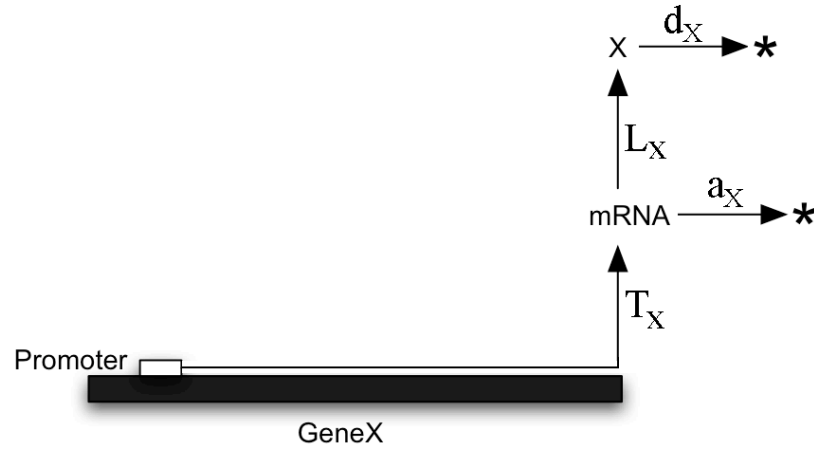


Figure B.1. A model for gene expression with transcription and translation.

scribed from the gene *GeneX* at a constant rate  $T_x$  and the protein  $X$  is translated from the mRNA at a constant rate  $L_x$ . Both mRNA and the protein decay at rates  $a_x$  and  $d_x$  respectively. As the average lifetime of a mRNA is  $1/a_x$  and proteins are made from it at rate  $L_x$ ,  $N_x = L_x/a_x$  denotes the number of proteins produced per mRNA, which is

referred to as the *burst size* of the gene *GeneX*. We denote by  $\mathbf{m}_x$  and  $\mathbf{x}$ , the number of molecules of the mRNA and protein X, respectively. As a continuous deterministic model based on chemical rate equations does not provide information about the stochastic fluctuation in the protein, we consider a stochastic formulation that treats births and deaths of the mRNA and the protein as probabilistic events. Given that  $\mathbf{x}(t) = x$  and  $\mathbf{m}_x(t) = m_x$ , the probabilities of the four reactions corresponding to births and deaths of the mRNA and the protein happening in the infinitesimal time interval  $(t, t + dt]$  are given by

$$\Pr\{\mathbf{x}(t + dt) = x, \mathbf{m}_x(t + dt) = m_x + 1\} = T_x dt \quad (\text{B.23a})$$

$$\Pr\{\mathbf{x}(t + dt) = x, \mathbf{m}_x(t + dt) = m_x - 1\} = a_x m_x dt \quad (\text{B.23b})$$

$$\Pr\{\mathbf{x}(t + dt) = x + 1, \mathbf{m}_x(t + dt) = m_x\} = L_x m_x dt \quad (\text{B.23c})$$

$$\Pr\{\mathbf{x}(t + dt) = x - 1, \mathbf{m}_x(t + dt) = m_x\} = d_x x dt. \quad (\text{B.23d})$$

We model the time evolution of the number of molecules  $\mathbf{x}$  and  $\mathbf{m}_x$  through a Stochastic Hybrid System (SHS), the state of which is  $\mathbf{y} = [\mathbf{m}_x, \mathbf{x}]^T$ . This SHS is characterized by trivial continuous dynamics

$$\dot{\mathbf{y}} = 0, \quad (\text{B.24})$$

four reset maps  $\phi_i(\mathbf{y})$

$$\mathbf{y} \mapsto \phi_1(\mathbf{y}) = \begin{bmatrix} \mathbf{m}_x + 1 \\ \mathbf{x} \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_2(\mathbf{y}) = \begin{bmatrix} \mathbf{m}_x - 1 \\ \mathbf{x} \end{bmatrix} \quad (\text{B.25a})$$

$$\mathbf{y} \mapsto \phi_3(\mathbf{y}) = \begin{bmatrix} \mathbf{m}_x \\ \mathbf{x} + 1 \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_4(\mathbf{y}) = \begin{bmatrix} \mathbf{m}_x \\ \mathbf{x} - 1 \end{bmatrix} \quad (\text{B.25b})$$

and corresponding transition intensities

$$\lambda_1(\mathbf{y}) = T_x, \quad \lambda_2(\mathbf{y}) = a_x \mathbf{m}_x, \quad \lambda_3(\mathbf{y}) = L_x \mathbf{m}_x, \quad \lambda_4(\mathbf{y}) = d_x \mathbf{x} \quad (\text{B.26})$$



corresponding to transcription, translation, mRNA and protein degradation. We now determine the time evolution of the first and second order moments of  $\mathbf{y}$ , i.e., the expected values  $\mathbf{E}[\mathbf{m}_x]$ ,  $\mathbf{E}[\mathbf{x}]$ ,  $\mathbf{E}[\mathbf{x}^2]$ ,  $\mathbf{E}[\mathbf{m}_x^2]$  and  $\mathbf{E}[\mathbf{m}_x\mathbf{x}]$ . Using the Dynkin's formula for the above SHS we have

$$\frac{d\mathbf{E}[\mathbf{m}_x]}{dt} = T_x - a_x\mathbf{E}[\mathbf{m}_x] \quad (\text{B.27a})$$

$$\frac{d\mathbf{E}[\mathbf{x}]}{dt} = L_x\mathbf{E}[\mathbf{m}_x] - d_x\mathbf{E}[\mathbf{x}] \quad (\text{B.27b})$$

$$\frac{d\mathbf{E}[\mathbf{m}_x^2]}{dt} = T_x + a_x\mathbf{E}[\mathbf{m}_x] + 2T_x\mathbf{E}[\mathbf{m}_x] - 2a_x\mathbf{E}[\mathbf{m}_x^2] \quad (\text{B.27c})$$

$$\frac{d\mathbf{E}[\mathbf{x}^2]}{dt} = L_x\mathbf{E}[\mathbf{m}_x] + d_x\mathbf{E}[\mathbf{x}] + 2L_x\mathbf{E}[\mathbf{m}_x\mathbf{x}] - 2d_x\mathbf{E}[\mathbf{x}^2] \quad (\text{B.27d})$$

$$\frac{d\mathbf{E}[\mathbf{m}_x\mathbf{x}]}{dt} = L_x\mathbf{E}[\mathbf{m}_x^2] + T_x\mathbf{E}[\mathbf{x}] - d_x\mathbf{E}[\mathbf{m}_x\mathbf{x}] - a_x\mathbf{E}[\mathbf{m}_x\mathbf{x}]. \quad (\text{B.27e})$$

The steady-state moments are then given by

$$\mathbf{m}_x^* := \lim_{t \rightarrow \infty} \mathbf{E}[\mathbf{m}_x(t)] = \frac{T_x}{a_x} \quad (\text{B.28a})$$

$$\mathbf{x}^* := \lim_{t \rightarrow \infty} \mathbf{E}[\mathbf{x}(t)] = \frac{L_x T_x}{d_x a_x} = \frac{N_x T_x}{d_x} \quad (\text{B.28b})$$

$$\mathbf{E}^*[\mathbf{m}_x^2] := \lim_{t \rightarrow \infty} \mathbf{E}[\mathbf{m}_x^2(t)] = \frac{a_x T_x + T_x^2}{a_x^2} \quad (\text{B.28c})$$

$$\mathbf{E}^*[\mathbf{x}^2] := \lim_{t \rightarrow \infty} \mathbf{E}[\mathbf{x}^2(t)] = \frac{L_x T_x}{d_x a_x} + \frac{L_x (d_x a_x L_x T_x + d_x L_x T_x^2 + a_x L_x T_x^2)}{d_x^2 a_x^2 (d_x + a_x)} \quad (\text{B.28d})$$

$$\mathbf{E}^*[\mathbf{m}_x\mathbf{x}] := \lim_{t \rightarrow \infty} \mathbf{E}[\mathbf{m}_x(t)\mathbf{x}(t)] = \frac{d_x a_x L_x T_x + d_x L_x T_x^2 + a_x L_x T_x^2}{d_x a_x^2 (d_x + a_x)}. \quad (\text{B.28e})$$

Using the above steady-state we obtain the following coefficient of variation of  $\mathbf{x}$

$$CV_{int-nr}^2 = \frac{1 + N_x + e}{\mathbf{x}^* (1 + e)} \quad (\text{B.29})$$

where  $e = d_x/a_x$  is the ratio of the protein and the mRNA degradation rate. As mentioned before  $e$  is generally small and in the limit where  $e$  goes to zero,  $CV_{int-nr}$  reduces to

$$CV_{int-nr}^2 = \frac{1 + N_x}{\mathbf{x}^*}. \quad (\text{B.30})$$

Comparing (B.30) with (4.11) we see that the noise in the protein population (when  $e \approx 0$ ) is the same as that obtained from the model presented in Section 4.1 with  $V_x^2 = N_x^2 + N_x$ .

## B.6 Auto-regulatory gene network with transcription and translation

We now put a feedback in the gene expression model introduced in Appendix B.5. We model the feedback by assuming that the transcription rate of the gene is given by  $g(\mathbf{x})$  when  $\mathbf{x}$  denotes the number of protein molecules in the cell. As done in Section 4.2.1 we use a linear approximation

$$g(\mathbf{x}) \approx g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \quad (\text{B.31})$$

where  $\mathbf{x}^*$  is the steady-state average number of protein molecules. The corresponding SHS for this auto-regulatory gene network has trivial dynamics and reset maps given by (B.25), and corresponding transition intensities

$$\lambda_1(\mathbf{y}) = g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*), \quad \lambda_2(\mathbf{y}) = a_x \mathbf{m}_x, \quad \lambda_3(\mathbf{y}) = L_x \mathbf{m}_x, \quad \lambda_4(\mathbf{y}) = d_x \mathbf{x}. \quad (\text{B.32})$$

Analysis similar to that done in Appendix B.5 shows that the steady-state coefficient of variation of  $\mathbf{x}$  is given by

$$CV_{int}^2 = \frac{d_x(1 + N_x)}{[g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)]N_x(1 + e)} + \frac{e}{(1 + e)\mathbf{x}^*} \quad (\text{B.33})$$

where  $e = d_x/a_x$  is the ratio of the protein and mRNA degradation rate. Notice that when  $e \approx 0$ , the noise in protein numbers reduces to

$$CV_{int}^2 = \frac{d_x(1 + N_x)}{[g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)]N_x} \quad (\text{B.34})$$

which is identical to that obtained in (4.20) with  $V_x^2 = N_x^2 + N_x$ .

## B.7 Minimum limit of noise suppression with mRNA dynamics

We now investigate the intrinsic noise in the protein population given by (B.33) for the special class transcriptional response given by

$$g(\mathbf{x}) = \frac{g_0}{1 + (a\mathbf{x})^M}. \quad (\text{B.35})$$

Substituting (B.35) in (B.33) we have

$$CV_{int} = \sqrt{\frac{[1 + (a\mathbf{x}^*)^M][e(1 + (a\mathbf{x}^*)^M + M(a\mathbf{x}^*)^M) + (1 + N_x)(1 + (a\mathbf{x}^*)^M)]}{(1 + N_x + e)(1 + (a\mathbf{x}^*)^M + M(a\mathbf{x}^*)^M)}} CV_{int-nr} \quad (\text{B.36})$$

where

$$CV_{int-nr}^2 = \frac{1 + N_x + e}{\mathbf{x}^*(1 + e)} \quad (\text{B.37})$$

is the intrinsic noise in the protein when there is no feedback (i.e.,  $a = 0$ ). When  $e \approx 0$ , the limit of noise suppression is given

$$\frac{CV_{int-min}}{CV_{int-nr}} = \frac{\sqrt{4M}}{M + 1} \quad (\text{B.38})$$

which is identical to that obtained in Section 4.4.1 (see equation (4.48)). Numerical analysis suggests that as we now increase  $e$  from zero (i.e., the mRNA half life is no longer small compared to protein half life) the limit of noise suppression decreases and is lower than what is predicted by (B.38). For example, for  $e = 1$  (i.e., the mRNA and protein have same half lives) numerical analysis shows that for a burst size of  $N_x = 10$  there can be at most a 4.7% reduction in intrinsic noise from  $CV_{int-nr}$  for  $M = 2$  where as for  $M = 4$  we can have a 17.4% reduction. These maximum reductions are slightly smaller compared to the 5.7% and 20% reduction in intrinsic noise for the same values of  $M$  (see Section 4.4.1) when  $e \approx 0$ .

## Appendix C

# Incorporating extrinsic noise in the cascade of genes

Consider an exogenous signal  $\mathbf{z}$  defined by the following birth-death process

$$\Pr\{\mathbf{z}(t + dt) = z + \mathbf{N}_z \mid \mathbf{z}(t) = z\} = K_z dt \quad (\text{C.1a})$$

$$\Pr\{\mathbf{z}(t + dt) = z - 1 \mid \mathbf{z}(t) = z\} = d_z z dt \quad (\text{C.1b})$$

where  $K_z$  and  $d_z$  represent the production and degradation rate of  $\mathbf{z}$ , respectively, and  $\mathbf{N}_z$  is a random variable with mean  $N_z$  and variance  $V_z^2$ . Let protein  $X_1$  be expressed from a gene *GeneX<sub>1</sub>* and given by the following birth-death process

$$\Pr\{\mathbf{x}_1(t + dt) = x_1 + \mathbf{N}_1 \mid \mathbf{x}_1(t) = x_1\} = K_1 dt \quad (\text{C.2a})$$

$$\Pr\{\mathbf{x}_1(t + dt) = x_1 - 1 \mid \mathbf{x}_1(t) = x_1\} = d_1 x_1 dt \quad (\text{C.2b})$$

where  $\mathbf{x}_1$  is the number of molecules of the protein  $X_1$ . Protein  $X_1$  activates a gene *GeneX<sub>2</sub>* to make protein  $X_2$  with the transcription rate of *GeneX<sub>2</sub>* given by  $\mathbf{z}\mathbf{x}_1$ . Our goal is to relate the noise in the protein  $X_2$  to the noise in protein  $X_1$  and the exogenous

signal  $\mathbf{z}$ . As before, we model the above process by a SHS with trivial dynamics

$$\dot{\mathbf{y}} = 0, \quad \mathbf{y} = [\mathbf{z}, \mathbf{x}_1, \mathbf{x}_2] \quad (\text{C.3})$$

with six reset maps

$$\mathbf{y} \mapsto \phi_1(\mathbf{y}) = \begin{bmatrix} \mathbf{z} + \mathbf{N}_z \\ \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_2(\mathbf{y}) = \begin{bmatrix} \mathbf{z} - 1 \\ \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad (\text{C.4a})$$

$$\mathbf{y} \mapsto \phi_3(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x}_1 + \mathbf{N}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_4(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x}_1 - 1 \\ \mathbf{x}_2 \end{bmatrix}, \quad (\text{C.4b})$$

$$\mathbf{y} \mapsto \phi_5(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x}_1 \\ \mathbf{x}_2 + \mathbf{N}_2 \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_6(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x}_1 \\ \mathbf{x}_2 - 1 \end{bmatrix}, \quad (\text{C.4c})$$

and corresponding transition intensities

$$\lambda_1(\mathbf{y}) = K_z, \quad \lambda_2(\mathbf{y}) = d_z \mathbf{z}, \quad \lambda_3(\mathbf{y}) = K_1, \quad \lambda_4(\mathbf{y}) = d_1 \mathbf{x}_1, \quad \lambda_5(\mathbf{y}) = \mathbf{z} \mathbf{x}_1, \quad \lambda_6(\mathbf{y}) = d_2 \mathbf{x}_2. \quad (\text{C.5})$$

We now construct a vector  $\mu$  with

$$\mu = [\mathbf{E}[\mathbf{z}], \mathbf{E}[\mathbf{x}_1], \mathbf{E}[\mathbf{x}_2], \mathbf{E}[\mathbf{z}^2], \mathbf{E}[\mathbf{x}_1^2], \mathbf{E}[\mathbf{x}_2^2], \mathbf{E}[\mathbf{z}\mathbf{x}_1], \mathbf{E}[\mathbf{z}\mathbf{x}_2], \mathbf{E}[\mathbf{x}_1\mathbf{x}_2], \mathbf{E}[\mathbf{z}\mathbf{x}_1\mathbf{x}_2]]^T. \quad (\text{C.6})$$

Using Dynkin's equations we have that the time derivative of  $\mu$  is given by

$$\dot{\mu} = \bar{a} + \bar{A}\mu + \bar{B}\bar{\mu}, \quad \bar{\mu} = [\mathbf{E}[\mathbf{z}^2\mathbf{x}_1], \mathbf{E}[\mathbf{z}\mathbf{x}_1^2], \mathbf{E}[\mathbf{z}^2\mathbf{x}_1^2]]^T \quad (\text{C.7})$$

for some vector  $\bar{a}$  and matrices  $\bar{A}$  and  $\bar{B}$ . Using the fact that the stochastic processes  $\mathbf{x}_1$  and  $\mathbf{z}$  are independent, we have

$$\bar{\mu} = [\mathbf{E}[\mathbf{z}^2]\mathbf{E}[\mathbf{x}_1], \mathbf{E}[\mathbf{z}]\mathbf{E}[\mathbf{x}_1^2], \mathbf{E}[\mathbf{z}^2]\mathbf{E}[\mathbf{x}_1^2]]^T \quad (\text{C.8})$$

which leads to a closed system of differential equations given by

$$\dot{\mu} = \bar{a} + \bar{A}\mu + \bar{B}\bar{\mu}, \quad \bar{\mu} = [\mathbf{E}[\mathbf{z}^2]\mathbf{E}[\mathbf{x}_1], \mathbf{E}[\mathbf{z}]\mathbf{E}[\mathbf{x}_1^2], \mathbf{E}[\mathbf{z}^2]\mathbf{E}[\mathbf{x}_1^2]]^T. \quad (\text{C.9})$$

Solving for the steady-state of (C.9) we have

$$CV_{X_2}^2 = \frac{N_2}{\mathbf{x}_2^*} + \xi CV_{X_1}^2 + \zeta CV_z^2 + \delta CV_z^2 CV_{X_1}^2 \quad (\text{C.10})$$

where

$$\xi = \frac{d_2}{d_1 + d_2} < 1, \quad \zeta = \frac{d_2}{d_z + d_2} < 1, \quad \delta = \frac{d_2}{d_z + d_2 + d_2} < 1, \quad (\text{C.11})$$

and  $CV_z, CV_{X_1}$  are the noise in the exogenous signal  $\mathbf{z}$  and protein  $X_1$ , respectively. The quantity  $\mathbf{x}_2^*$  denotes the steady-state average number of molecules of protein  $X_2$ .