Model order reduction for stochastic models of biomolecular systems with time-scale separation

by

Narmada Kumari Herath

B.Eng., University of Sheffield (2011)
S.M., Massachusetts Institute of Technology (2013)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

© Massachusetts Institute of Technology 2018. All rights reserved.

Signature redacted

Author .................................................................

Department of Electrical Engineering and Computer Science

Certified by .............. ...........................................

Domitilla Del Vecchio
Professor of Mechanical Engineering
Thesis Supervisor

Accepted by .................... ......................................

Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students
Model order reduction for stochastic models of biomolecular systems with time-scale separation

by

Narmada Kumari Herath

Submitted to the Department of Electrical Engineering and Computer Science on May 23, 2018, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

Biomolecular systems often involve reactions that take place on different time-scales, giving rise to 'slow' and 'fast' system variables. This property is widely used in the analysis of systems to obtain dynamical models with reduced dimensions. In deterministic systems, methods to obtain such reduced-order models are well defined by the singular perturbation or averaging techniques. However, model reduction of stochastic systems remains an ongoing area of research. In particular, existing model reduction methods for stochastic models of biomolecular systems lack rigorous error quantifications between the full and reduced dynamics. Furthermore, they only provide approximations for the slow variable dynamics, making the application of such methods to biomolecular systems difficult since the variables of interest are typically mixed (i.e., they encompass both fast and slow variables).

In this thesis, we consider biomolecular systems modeled using the chemical Langevin equation (CLE) and the Linear Noise Approximation (LNA). Specifically, we consider biomolecular systems with linear propensity functions modeled by the CLE and systems with arbitrary propensity functions modeled by the LNA. For these systems, we obtain reduced-order models that approximate both the slow and fast variables under time-scale separation conditions. In particular, with suitable assumptions, we prove that the moments of the reduced-order models converge to those of the full systems as the time-scale separation becomes large. Our results further provide a rigorous justification for the accuracy of the stochastic total quasi-steady state approximation (tQSSA).

We then consider two applications of these reduced-order models. In the first application, we analyze the trade-offs between modularity and signal noise in biomolecular networks. In the second application, we consider the application of the reduced-order LNA developed in this work to obtain reduced-order stochastic models for gene-regulatory networks.

Thesis Supervisor: Domitilla Del Vecchio
Title: Professor of Mechanical Engineering
Acknowledgments

First and foremost, I would like to thank my advisor Professor Domitilla Del Vecchio for her invaluable guidance and support throughout my graduate studies. Her passion and enthusiasm for research has been deeply inspiring and this thesis would not have been possible without her consistent advice, patience and feedback. I feel very fortunate to have been a part of her research group.

I also thank the members of my thesis committee, Professor George Verghese and Professor Devavrat Shah, for their valuable feedback that broadened my view on many different aspects of this work.

I am grateful to the past and present members of the Del Vecchio group for their insights and feedback during my time at MIT. I would especially like to thank Simone Bruno for his contributions in the final stage of this work, Mohammad Naghnaeian for the many stimulating discussions and valuable suggestions over the past couple of years, Abdullah Hamadeh for his guidance and the contribution to the initial stages of this work, Daniel Hochener and Andrea Boccia for their insights and comments on working with stochastic differential equations, Yili Qian and Rushina Shah for always answering my ‘quick’ questions, Heejin Ahn and Penny Chen for their friendship and all other members of the group for the friendly and helpful work environment they have created. I would also like to thank Joe Gaken for his prompt assistance in every administrative need.

I would also like to express my sincere gratitude towards the professors and mentors that I met during my undergraduate studies at the University of Sheffield, in particular, Professor Visakan Kadirkamanathan, Dr. Roderich Gross and Ms. Linda Gray. Without their support and encouragement in the final year of my undergraduate studies, coming to MIT would have remained a dream.

I am deeply grateful to all the friends I have met during my time here. Particularly, I would like to thank Prabhani, Oshani, Sachi and Garthee for their wonderful company when I first came to MIT and helping me get settled into the student life in USA. I want to thank Charith, Maleen, Sandamali, Thulith, Nawodya, Chamithri and
Tharanga for the never-ending laughter and fun during our many game nights, hikes and movie outings. Savithru, my deepest thanks goes to you for being a constant source of encouragement, strength and motivation throughout this journey. Your unwavering support has always inspired me to do my best and I am grateful for all memories we have shared over the years.

Lastly, I would like to thank my parents for their support and encouragement, and especially my sister for being my very first role model and inspiring me to pursue a higher education.
## Contents

1 Introduction 15
   1.1 Literature review on model order reduction of CLE and LNA 19
   1.2 Thesis organization 21

2 Background 23
   2.1 Stochastic models of biomolecular systems 23
      2.1.1 Chemical Master Equation (CME) 24
      2.1.2 Chemical Langevin Equation (CLE) 25
      2.1.3 Linear Noise Approximation (LNA) 27
   2.2 Gene-regulatory networks 28

3 Reduced-order model for multi-scale chemical Langevin equation with linear propensity functions 31
   3.1 Introduction 31
   3.2 System model 32
   3.3 Results 34
      3.3.1 Reduced-order model 34
      3.3.2 Preliminary results 35
      3.3.3 Main results 37
   3.4 Academic example 41
   3.5 Conclusion 44

4 Reduced-order model for multi-scale Linear Noise Approximation:
The stochastic tQSSA$^+$

4.1 Introduction ............................................. 45
4.2 System model ............................................. 46
4.3 Results .................................................. 49
  4.3.1 Reduced-order model ................................. 49
  4.3.2 Preliminary results ................................... 50
  4.3.3 Main results .......................................... 52
4.4 Validity of LNA models derived using stochastic tQSSA ....... 56
4.5 Conclusion ............................................... 59

5 Interplay between modularity and signal noise in biomolecular networks ............................................. 61
  5.1 Introduction .............................................. 61
  5.2 Modularity in gene-regulatory networks ................. 61
  5.3 Trade-off between retroactivity and noise in a cascade of transcriptional components ................................. 63
    5.3.1 Retroactivity error ................................... 69
    5.3.2 Noise error .......................................... 70
  5.4 Trade-off between resource competition and noise .......... 73
    5.4.1 Error in Y due to resource competition ............... 77
    5.4.2 Noise error in G .................................... 77
  5.5 Noise buffering through decoy sites ........................ 78
    5.5.1 Noise analysis ....................................... 83
  5.6 Insulation through phosphorylation cycles ................. 85
    5.6.1 Noise analysis ....................................... 90
  5.7 Conclusion ............................................... 94

6 Reduced-order LNA models for gene-regulatory networks .... 97
  6.1 Introduction .............................................. 97
  6.2 Network structure and problem description ................. 98
    6.2.1 Problem description .................................. 100
6.3 Results ........................................................................................................... 101
  6.3.1 Hill function model ................................................................................... 101
  6.3.2 Moments of the stochastic tQSSA\textsuperscript{+} model ......................... 102
6.4 Discussion and future work ........................................................................... 105
6.5 Model derivations ......................................................................................... 109
  6.5.1 Macroscopic reaction rate equations ....................................................... 109
  6.5.2 Definitions ................................................................................................ 110
  6.5.3 Hill function model .................................................................................. 111
  6.5.4 Stochastic tQSSA\textsuperscript{+} model .......................................................... 112

7 Conclusion ........................................................................................................ 125
  7.1 Contributions ............................................................................................... 125
  7.2 Future directions ......................................................................................... 127

A Proofs pertaining to Chapter 3 ........................................................................ 131
  A.1 Proof of Claim 3.1 ....................................................................................... 131
  A.2 Proof of Claim 3.2 ....................................................................................... 134
  A.3 Proof of Claim 3.3 ....................................................................................... 135
  A.4 Proof of Theorem 3.1 .................................................................................. 138

B Proofs pertaining to Chapter 4 ........................................................................ 143
  B.1 Proof of Claim 4.1 ....................................................................................... 143
  B.2 Proof of Claim 4.2 ....................................................................................... 145
  B.3 Proof of Claim 4.3 ....................................................................................... 146
  B.4 Proof of Claim 4.4 ....................................................................................... 147
  B.5 Proof of Theorem 4.1 .................................................................................. 149
  B.6 Proof for convergence of reduced and the original moments .................. 156

C Derivations and simulation results pertaining to Chapter 5 ......................... 159
  C.1 Derivation of the moment dynamics of the mixed variable .................... 159
  C.2 Derivations for Section 5.3 ....................................................................... 161
  C.3 Derivations for Section 5.6 ....................................................................... 162
C.4 Simulations of the full and reduced systems .................. 164
C.4.1 Simulation results for Section 5.5 ......................... 164
C.4.2 Simulation results for Section 5.6 ......................... 166

D Derivations pertaining to Chapter 6 .......................... 167
D.1 Stability of the slow manifold ................................. 167
D.2 Matrix inverse ................................................. 168
D.3 System models for incoherent feedforward loop ............ 171
  D.3.1 Hill function model ......................................... 172
  D.3.2 Stochastic tQSSA+ model ................................. 173
List of Figures

2-1 Schematic of a gene-regulatory network. Node $i$ is an input to node $j$, which is in turn an input to node $k$. ................................. 29

3-1 Setting $\epsilon = 0$ in the moment dynamics of the full system yields the moment dynamics of the reduced system. ................................. 38

3-2 Power spectrum of $z$ (top). Frequency response from $z$ to $x$ (centre). Power spectrum of $x$ (bottom). The parameters used are $a_1 = 10$, $a_2 = 2$ and $v_1 = 1$, $v_2 = 10$. ......................................................... 41

3-3 Moments of the full and reduced systems. (a) Second moments of the slow variable. (b) Third moments of the slow variable. (c) Second moments of the fast variable. ................................. 42

4-1 Setting $\epsilon = 0$ in the moment dynamics of the full system yields the moment dynamics of the reduced system. ................................. 53

4-2 Schematic diagram illustrating the model reduction approach given by Theorem 4.1. The $O(\epsilon)$-closeness of the moments of the original species concentrations and the moments obtained via the reduced system is proved in Appendix B.6. ................................. 54

4-3 Schematic diagram illustrating the relationship between the singular perturbation approach and tQSSA approach. ................................. 58

5-1 Protein $X$ acts as an input to the upstream transcriptional component, which produces the output protein $Y$. The downstream transcriptional components takes protein $Y$ as an input and produces protein $G$. ................................. 63
5-2 Nominal and perturbed signals. $G$ is obtained by simulating system (5.1) using the Gillespie algorithm [21]. The parameter values are $X = 2 + 1.5 \sin(\omega t) \text{nM}$, $\alpha_1 = 1 \text{nM}^{-1}s^{-1}$, $\alpha_2 = 20s^{-1}$, $\alpha_3 = 1 \text{nM}^{-1}s^{-1}$, $\alpha_4 = 100s^{-1}$, $\beta_1 = 0.01s^{-1}$, $\beta_2 = 0.1s^{-1}$, $\delta_1 = \delta_2 = 0.01s^{-1}$, $p_{T0} = 100 \text{nM}$ and $\omega = 0.002 \text{ rad/s}$.

5-3 The signal 'r' denotes the retroactivity to the upstream system. System 1 represents the nominal system in the absence of any perturbations. System 2 represents the system is perturbed only with retroactivity. System 3 represents the system perturbed with both retroactivity and intrinsic noise. $\Gamma_Y$ encapsulates the noise in the upstream component given by $\Gamma_i$ for $i = 1, ..., 4$ and $\Gamma_G$ encapsulates the noise in the downstream component given by $\Gamma_i$ for $i = 5, ..., 8$.

5-4 Trade-off between retroactivity and noise in signal $G$ for $p_T$ in the range 1:1000 nM (obtained using the equations (5.14) and (5.21)). The parameter values are as in Figure 4.

5-5 The proteins Y and G are expressed constitutively, in which X is a resource shared during the process of protein production.

5-6 Concentrations of Y and G obtained by simulating system (5.25) using the Gillespie algorithm [21]. The parameter values are $X = 10 \text{nM}$, $\beta_1 = \beta_2 = 0.1s^{-1}$, $\delta_1 = \delta_2 = 0.01s^{-1}$, $p_{T1} = 200 \text{nM}$, $k_{d1} = k_{d2} = 100 \text{nM}$.

5-7 Trade-off between reducing noise in G and increasing the error due to resource competition in Y for $p_T$ in the range 50:1000 nM (obtained using the equations (5.29) and (5.30)). The parameter values are $X = 10 \text{nM}$, $k_{d1} = 100 \text{nM}$, $\beta_2 = 0.01s^{-1}$, $\delta_2 = 0.01s^{-1}$, $p_{T1} = 100 \text{nM}$.

5-8 Transcription factor X regulates the production of protein G, while also binding to non-regulatory binding site $p_1$. 
5-9 Increasing decoy sites $p_1$ shifts the noise content of signal X to high frequency. Simulation performed with Gillespie algorithm. The parameters used are $k = 1\text{nM} \cdot \text{sec}^{-1}$, $\delta_1 = 0.1\text{sec}^{-1}$, $\delta_1 = 0.1\text{sec}^{-1}$, $k_{on1} = 1\text{sec}^{-1}$, $k_{off1} = 100\text{nM} \cdot \text{sec}^{-1}$, $k_{on2} = 1\text{sec}^{-1}$, $k_{off2} = 100\text{nM} \cdot \text{sec}^{-1}$, $p_{T2} = 1\text{nM}$, $\beta = 10\text{sec}^{-1}$.

5-10 Protein X is phosphorylated by kinase Z and dephosphorylated by phosphatase Y. Phosphorylated protein $X^*$ binds to the downstream promoter.

5-11 Maximum coefficient of variation for proteins G and $X^*$. The phosphorylation and dephosphorylation rate constants were approximated from Chen et al [10] as $k_1 = 0.0005 \text{nm} \cdot \text{min}^{-1}$ and $k_2 = 0.0005 \text{nm} \cdot \text{min}^{-1}$. The remaining parameters were chosen as $\beta = 0.5 \text{min}^{-1}$, $\delta = 0.01 \text{min}^{-1}$, $p_T = 5 \text{nM}$, $k_d = 10 \text{nM}$ to be in the nominal range for bacterial cells [27].

6-1 (a) Gene-regulatory network. (b) Promoter-TF complexes for a node with two inputs, where each input binds independently as considered in this work.

6-2 Protein $Y_1$ activates the production of protein $Y_2$ and $Y_3$, while $Y_2$ represses the production of protein $Y_3$.

6-3 Sample variances computed using the Hill function based LNA model and the stochastic tQSSA$^+$ based LNA model. Sample variances are computed using 150,000 simulation runs. The parameter values are $\pi_{10} = 0.01 \text{min}^{-1}$, $\pi_{21} = 1 \text{min}^{-1}$, $\pi_{31} = 100 \text{min}^{-1}$, $\delta_1 = \delta_2 = \delta_3 = 0.01 \text{min}^{-1}$, $p_{T1} = 1 \text{nM}$, $p_{T2} = 1 \text{nM}$, $p_{T3} = 1 \text{nM}$. The dissociation constants for the top figure are $k_{d31} = 10 \text{nM}$, $k_{d32} = 5 \text{nM}$, and for the bottom figure are $k_{d31} = 1000 \text{nM}$, $k_{d32} = 100 \text{nM}$. The chemical reactions and the system models are given in Appendix D.3.
6-4 Sample covariances computed using the Hill function based LNA model and the stochastic tQSSA+ based LNA model. Sample variances are computed using 150,000 simulation runs. The parameter values are $\pi_{10} = 0.01 \text{ min}^{-1}$, $\pi_{21} = 1 \text{ min}^{-1}$, $\pi_{31} = 100 \text{ min}^{-1}$, $\delta_1 = \delta_2 = \delta_3 = 0.01 \text{ min}^{-1}$, $p_{T1} = 1 \text{ nM}$, $p_{T2} = 1 \text{ nM}$, $p_{T3} = 1 \text{ nM}$. The dissociation constants for the top figure are $k_{d31} = 10 \text{ nM}$, $k_{d32} = 5 \text{ nM}$, and for the bottom figure are $k_{d31} = 1000 \text{ nM}$, $k_{d32} = 100 \text{ nM}$. The chemical reactions and the system models are given in Appendix D.

C-1 Errors in the first and second order moments. Average of 100,000 runs. The parameters used are $k = 10$, $\delta_1 = 0.1$, $\delta_1 = 1$, $k_{d1} = 1000$, $k_{d2} = 1000$, $P_{11} = 1000$, $P_{22} = 1000$, $\beta = 1$, $y(0) = 70$, $G(0) = 60$, $C_1(0) = 20$ and $C_2(0) = 20$.

C-2 Moments of the original and reduced systems. (a) Second moments of $\psi_v$ and $\bar{\psi}_v$. (b) Second moments of $\psi_g$ and $\bar{\psi}_g$. (c) Second moments of $\psi_c$ and $\bar{\psi}_c$. 

14
Chapter 1

Introduction

Many biochemical processes involve reactions that occur on different time-scales. For example, in bacterial cells, the binding of transcription factor to DNA takes place on the time-scale of seconds, while protein production and dilution are on the order of hours [2]. Such a separation in time-scales allows the system variables to be separated into slow and fast groups, and this property can be exploited to reduce the complexity of dynamical models. In particular, for deterministic systems, the quasi-steady state approximation gives a reduced-order model for the slow variables, assuming that the fast variables rapidly reach a steady state [7, 30]. In the mathematical literature, a system of ordinary differential equations (ODEs) with multiple time-scales is represented as slow and fast subsystems by using a small parameter $\epsilon$ to capture the separation in time-scales. The mathematical treatment of such systems is given by two main methods: singular perturbation and averaging [45, 66]. The singular perturbation approach, formalized by Tikhonov’s theorem, involves setting $\epsilon = 0$ in the system dynamics to obtain an algebraic equation that approximates the fast variables, which is in turn used to derive an approximation for the slow variable dynamics [89, 45, 50]. In the averaging method, a reduced-order model for the slow variables is obtained by the elimination of the fast dynamics via integration of the system functions [66].

As opposed to deterministic models, employing time-scale separation for model order reduction remains an ongoing area of research for stochastic models of biological
systems. Yet, obtaining reduced descriptions of stochastic dynamics is even more critical than for deterministic dynamics in order to increase the speed of simulation and aid analytical studies. Furthermore, accurate reduced-order models are important for precise parameter estimations [11].

The most prominent model used to capture the stochasticity in biological systems is the Chemical Master Equation (CME), which considers the species counts as a set of discrete random variables and describes the time evolution of their probability distributions using a set of ordinary differential equations [19, 90]. However, analyzing the chemical Master equation directly proves to be a challenge due to the lack of analytical tools to analyze its behavior. Therefore, several approximations of the Master equation have been developed, which provide good descriptions of the system dynamics under certain assumptions. The chemical Langevin equation (CLE) is one such approximation, where the dynamics of the chemical species are described as a set of stochastic differential equations [22]. The Fokker-Plank equation is another method equivalent to the CLE, which considers the species counts as continuous variables and provides a description of the time evolution of their probability density functions as a partial differential equation [19]. The Linear Noise Approximation (LNA) is another approximation, where the system dynamics are portrayed as stochastic fluctuations about a deterministic trajectory, assuming that the system volume is sufficiently large such that the fluctuations are small relative to the average species counts [90, 17]. The stochastic fluctuations in the LNA can be modeled using stochastic differential equations [65] or partial differential equations [90].

There have been several works that obtain reduced-order representations of the chemical Master equation [67, 58, 26, 38, 39, 76, 20, 9, 70, 47, 44, 48, 54, 75, 49]. Among these, a common approach used to reduce the complexity in the simulations of the CME is to approximate the fast variables by their deterministic quasi-steady state expressions [70]. This method is termed stochastic QSSA [3, 47] in literature. However, validity of this method still remains under investigation [73, 47, 88, 48, 49]. Elimination of fast variables in Fokker-Planck equations has also been considered in [86]. More recently, model reduction methods for chemical Langevin equation
and Linear Noise Approximation have also appeared [52, 13, 65, 81]. In addition to the above methods in systems/mathematical biology, mathematical literature on stochastic differential equations also allows for model reduction of chemical Langevin equations by using averaging methods [46, 66]. Recently, averaging methods have been applied in the analysis of systems modeled by the chemical Langevin equation [92].

Due to the intractability of the chemical Master equation, it is more desirable to use approximate models such as the chemical Langevin equation or Linear Noise Approximation, for purposes of theoretical analysis. However, the works in [52, 13, 65] on obtaining reduced-order systems for these models do not provide an error quantification between the full and reduced systems. The work by Sootla and Anderson in [81] obtains a reduced-order Linear Noise Approximation and provides an error quantification, but it is assumed that the diffusion coefficients of the LNA are Lipschitz continuous, which is not guaranteed in general. Mathematical literature for SDEs include several singular perturbation approaches for stochastic systems. However, these methods cannot be applied for the class of SDEs that arise due to the CLE and the LNA. Averaging methods in mathematical literature proves convergence in distribution between the full and reduced models, however, this method requires the integration of the system vector field which may be undesirable when the system is nonlinear or is of high dimension. Furthermore, the above methods do not provide an approximation for the fast variable dynamics. However, in biomolecular systems, it is typically important to approximate both slow and fast variables as chemical species often participate in both slow and fast reactions, and hence the corresponding concentrations are neither slow nor fast variables, but instead are mixed variables. In these systems, a coordinate transformation can be employed to take the system to standard singular perturbation form [13], in which fast and slow variables may not directly correspond to the physical variables of interest. Thus, we often require approximations for both slow and fast variables to analyze the stochastic properties of the variable of interest.

In this work, we consider biomolecular systems in multiple time-scales modeled
using the Chemical Langevin Equations and Linear Noise Approximation. We provide a model reduction framework that consists reduced-order approximations for both slow and fast variables, particularly, for CLE models with linear propensity functions and for LNA models. We then provide a mathematical error quantification between the full and reduced-order models in terms of the singular perturbation parameter. In particular, we use the moments of the full and reduced systems as a measure of error. We show that, under suitable assumptions, the moments of the reduced system are within an $O(\epsilon)$-neighborhood of the moments of the full system, and therefore the reduced system is a good approximation of the full system. Our results for the LNA in particular, provide a rigorous justification for reduced-order models derived using the stochastic tQSSA approach in literature.

We then consider two applications. In the first application, we analyze the interplay between modularity and signal noise in biomolecular network. Modularity allows the interconnection biomolecular components without causing a perturbation in the signal transmitted between the components. However, conditions for modularity often imply higher noise in the signals. We consider several network motifs that commonly occur in both natural and synthetic systems and use reduced-order CLE models and LNA models to quantify the stochastic effects of these systems under conditions that guarantee modularity of the network components. In the second application, we consider the problem of obtaining reduced-order stochastic models for gene-regulatory networks. In particular, we use the reduced-order LNA developed in this work to obtain a set of reduced moment dynamics for gene-regulatory networks, which can then be used to investigate the validity of commonly used stochastic quasi-steady state models in the literature.
1.1 Literature review on model order reduction of CLE and LNA

There have been several prior attempts at obtaining a reduced-order model for biomolecular systems modeled using chemical Langevin equations and Linear Noise Approximation. The work by Cantou-Carrere et al. considers chemical Langevin equations with multiple time-scales and obtains an approximation of the slow variable dynamics in the form of a Fokker-Plank equation, by assuming that the fast dynamics relaxes to a pseudo-stationary probability density while the slow dynamics remain constant [13]. The work by Pahlajani, et al. considers the Linear Noise Approximation where the stochastic fluctuations are modeled using stochastic differential equations [65]. They perform a singular perturbation analysis of the Backward Kolmogorov PDEs for the stochastic fluctuations in the LNA to derive a reduced order stochastic differential equation that approximates the slow dynamics of the full system. In [87, 88], Thomas et. al, derive a reduced order model for the slow variable by adiabatic elimination of the fast variables. In these methods, the error between the full system dynamics and the reduced dynamics is analyzed numerically and is not analytically quantified. Moreover, the reduced-order models only approximate the slow variable dynamics and do not provide an approximation for the fast dynamics. The work by Sootla and Anderson in [80] gives a projection-based model order reduction method to approximate the slow variable dynamics of systems modeled by the Linear Noise Approximation. The authors extend this work in [81], where they also provide an error quantification in the mean square sense for the reduced order model derived in [87] under quasi-steady state assumptions. However, to provide an error bound the authors explicitly use the Lipschitz continuity of the diffusion term, which is not Lipschitz continuous in general.

In addition to the above work, mathematical literature offers several methods for approaching singularly perturbed stochastic differential equations. Kabanov and Pergamenshchikov provide a stochastic version of the Tikhonov’s theorem for systems where the diffusion coefficient of the fast variable is of $o(\sqrt{e}/\sqrt{\ln(e)})$ [43]. Their
results show that when the time-scale separation becomes large ($\epsilon$ becomes small), the reduced system converges in probability to the full system, under the standard singular perturbation assumption that the slow manifold is exponentially stable. They also discuss another class of systems where the diffusion coefficient for the fast variable is $O(\sqrt{\epsilon})$. For this type of systems, the fast variable may be oscillatory and it may not converge in probability to the slow manifold. Another study by Berglund and Gentz uses a sample-path approach to find the probability of the solution being concentrated around a neighborhood of the slow manifold [5]. However, their analysis predicts that as $\epsilon$ decreases, the probability of the trajectory of the fast variable escaping a neighborhood of the slow manifold increases for the case where the diffusion coefficient is of $O(\sqrt{\epsilon})$. They also find an approximation for the slow variable, but this approximation is defined for the time-interval that the fast variable is within a neighborhood of the slow manifold. In [50] Kokotovic et al. developed a singular perturbation approach for linear stochastic systems in which the diffusion coefficient is a constant term. The analysis also includes systems where the diffusion coefficient is scaled by the singular perturbation parameter $\epsilon$. In both cases, they obtain a reduced system that converges to the full system in the mean squared sense. In [85], Tang and Basar approached the problem of singularly perturbed stochastic systems using the notion of stochastic input-to-state stability. They obtain stability results for the full system under the assumptions that the reduced fast and slow subsystems are input-to-state stable.

Another method in mathematical literature that can be used to analyze multi-scale stochastic differential equations is the averaging principle. The averaging principle by Khasminskii considers multiple scale stochastic differential equations where the diffusion term is of the order $\sqrt{\epsilon}$, and provides a reduced order model to which the full slow variable dynamics converge in distribution [46]. However, this method requires the explicit integration of the drift and diffusion coefficients of the full system, which may be undesirable if the system is nonlinear or of high dimension.
1.2 Thesis organization

In Chapter 2, we provide an introduction to the basic concepts used throughout this thesis. We first introduce the stochastic models that can be used to model the randomness in biomolecular systems. We then provide an introduction to gene-regulatory networks.

In Chapter 3, we consider the model order reduction of chemical Langevin equations with linear propensity functions and present a reduced-order model that approximates both slow and fast variable dynamics when the time-scale separation is large. We show that, on a finite time interval, moments of all orders of the slow variables for the full system become closer to those of the reduced-order model as time-scale separation is increased. A similar result holds for the first and second moments of the fast variable approximation.

In Chapter 4, we consider the model order reduction of Linear Noise Approximation. Similar to the previous chapter, we obtain a reduced-order LNA that approximates both slow and fast variables in the system, under time-scale separation conditions. In particular, we prove that the first and second moments of this reduced-order model converge to those of the full system as the time-scale separation becomes large.

In Chapter 5, we use the model reduction results developed in this work to analyze the interplay between modularity and signal noise in several gene-regulatory network motifs.

In Chapter 6, we use the reduced-order LNA model developed in Chapter 4 to derive a reduced-order moment description for gene-regulatory network motifs. We use these moment dynamics to investigate the validity of the LNA models for gene-regulatory networks derived using the stochastic quasi-steady state approximation.
Chapter 2

Background

In this chapter we provide an introduction to the basic concepts that are fundamental to the work in this thesis. First, we provide an introduction to the stochastic models used to model biomolecular systems. Next, we introduce gene-regulatory networks.

2.1 Stochastic models of biomolecular systems

Consider a biomolecular system with $n$ species $Y_1, \ldots, Y_n$, in a well-stirred volume $\Omega$, that chemically interact through $m$ reactions $R_1, \ldots, R_m$. Each reaction $R_i$ can be represented as:

$$p_{i1}Y_1 + \ldots + p_{in}Y_n \xrightarrow{k_i} r_{i1}Y_1 + \ldots + r_{in}Y_n, \quad i = 1, \ldots, m, \quad (2.1)$$

where $k_i$ denotes the rate constant of reaction $i$ and $r_{il} - p_{il}$ is the change in the number of molecules of $Y_l$ due to the reaction $i$.

The state of the system at a given time $t$ can be represented by the vector $y(t) = [y_1, \ldots, y_n]^T$, in which each component $y_i$ represents the molecular count for each species as a discrete random variable. Then, a reaction channel $R_i$ can be completely specified by the following quantities [22]:

(a) propensity function or the microscopic reaction rate $a_i(y)$, which defines the probability that the reaction will occur in a small time step $dt$:
\( a_i(y)dt \equiv \) probability that one reaction \( i \) will take place within \( \Omega \) in an infinitesimal time step \( dt \), given that the system is in state \( y \).

(b) **stoichiometry vector** \( q_i = [q_{i1}, \ldots, q_{in}] \), which defines the change in the number of species due to the firing of reaction \( R_i \):

For reaction \( R_i \) in the form (2.1): \( q_{ji} = r_{ji} - p_{ji} \)

The propensity function takes the form \( a_i(y) = k_i h_i(y) \) where \( k_i \) is the number of distinct combinations the reactant molecules of \( R_i \) available in state \( y \) [22]. For example, a unimolecular reaction of the form

\[ k_i y_i \rightarrow . \]

will have a propensity function of the form \( k_i y_i \). For a bimolecular reaction of the form

\[ k_i y_i y_k \rightarrow . \]

we have \( a_i(y) = \frac{k_i}{\Omega} y_i y_k \), since the probability of two molecules reacting together in a volume \( \Omega \) is inversely proportional to the volume. In the case where molecules of the same type reacting together

\[ k_i y_i y_i \rightarrow . \]

the propensity function is in the form \( a_i(y) = \frac{k_i}{\Omega} y_i(y_i - 1)/2 \), where the \( y_i(y_i - 1)/2 \) is the total number of ways that two molecules can be selected out of \( y_i \) molecules [14].

### 2.1.1 Chemical Master Equation (CME)

The CME considers the state of the system \( y(t) \) at time \( t \) and describes how the probability \( P(y, t) \) that the system is in state \( y \) at time \( t \) evolves over time. Using the propensity function and the stoichiometry matrix, one can compute the probability that the system is in state \( y \) at time \( t + dt \), given the state at time \( t \) as:

\[
P(y, t + dt) = P(y, t) \prod_{i=1}^{m} (1 - a_i(y)dt) + \sum_{i=1}^{m} P(y - q_i)a_i(y - q_i)dt,
\]
\[ P(y, t) + \sum_{i=1}^{m} [a_i(y - q_i)P(y - q_i, t) - a_i(y)P(y, t)] + O(dt^2), \]

where \( O(dt^2) \) represent the higher order terms in \( dt \).

As \( dt \) is an infinitely small time increment, taking the limit of \( dt \to 0 \) yields the

**Chemical Master Equation:**

\[ \frac{dP(y, t)}{dt} = \sum_{i=1}^{m} [a_i(y - q_i)P(y - q_i) - a_i(y)P(y, t)]. \quad (2.2) \]

Although, the CME provides a complete description of the system dynamics over time, it is challenging analyze it directly due to its large system size. Therefore, we next look at several approximate models that has been derived from the chemical Master equation under certain assumptions.

### 2.1.2 Chemical Langevin Equation (CLE)

The Chemical Langevin Equation is an approximation to the CME, derived under assumptions that are satisfied when the molecular counts and the system volume are sufficiently large. Introduced by D. T. Gillespie in [22], the derivation of the chemical Langevin equation starts by taking an approximation for the state of the system \( y(t + dt) \) in the form

\[ y_i(t + dt) = y_i(t) + \sum_{j=1}^{m} K_j(y, dt)q_{ji}, \quad (2.3) \]

where \( K_j(y, t) \) is defined to be a random variable that represents the number of reactions that occur in the time interval \( dt \). Then, to determine an approximation for the functions \( K_j(y, t) \), Gillespie considers the following assumptions:

(i) Require \( dt \) to be small enough such that the change in state during the time interval \([t, t + dt]\) will be very small and thus none of the propensity functions will change its value appreciably, i.e.

\[ a_j(y(t')) = a_j(y(t)), \quad t' \in [t, t + dt]. \]


This allows the occurrence of reactions to be independent of each other and the functions $K_j(y, dt)$ to be represented by independent Poisson random variables, i.e. $K_j(y, dt) = P_j(a_j(y), dt)$, where $P_j(a_j(y), dt)$ denotes a Poisson random variable with parameter $a_j(y)$.

(ii) Require $dt$ to be large enough such that the expected number of occurrences for each reaction $R_j$ in the time interval $[t, t + dt]$ is larger than 1. i.e.

$$E[P_j(a_j(y), dt)] = a_j(y)dt > 1.$$ 

This condition allows the Poisson random variable to be approximated by a normal random variable with the same mean the variance, which leads to

$$y_i(t + dt) = y_i(t) + \sum_{j=1}^{m} q_{ji}N_j(a_j(y)dt, a_j(y)dt),$$

where $N_j(m, \sigma^2)$ denotes a normal random variable with mean $m$ and variance $\sigma^2$.

As described in [22], the approximation of the integer Poisson random variable by the real normal random variables converts the molecular populations $y_i$ from discrete integer variables to continuous real variables. Next, using linear combination theory for normal random variables, giving $N_j(m, \sigma^2) = m + \sigma N_j(m, 1)$, the equation (2.4) can be written as

$$y_i(t + dt) = y_i(t) + \sum_{j=1}^{m} q_{ji}a_j(y)dt + \sum_{j=1}^{m} q_{ji}N_j(0, 1)(dt)^{1/2},$$

which is known as the chemical Langevin equation (CLE) [22]. Rearranging the terms and taking $dt \to 0$ leads to the white-noise form of the chemical Langevin equation:

$$\frac{dy_i(t)}{dt} = \sum_{j=1}^{m} q_{ji}a_j(y) + \sum_{j=1}^{m} q_{ji}\sqrt{a_j(y)} \Gamma_j,$$

where $\Gamma_j$ represents a Gaussian white noise process, that satisfies $E[\Gamma_j(t)\Gamma_j'(t')] = \delta(j, j')\delta(t - t')$ where the first delta function is Kronecker's and the second is Dirac's
The validity of this approximation depends on the existence of an infinitesimal time interval $dt$ that satisfies both assumptions (i) and (ii), which however, run counter to each other. In [24], D. T. Gillespie proves that it is always possible to find such a $dt$ if the system is sufficiently close to the thermodynamic limit, which is defined as the limit in which the molecular counts $y$ and system volume $\Omega$ becomes infinitely large in a way that the concentration $v = y/\Omega$ remains constant.

In addition to equation (2.5), we often represent the state of the system in terms of the vector of concentrations $v(t) = [v_1(t), \ldots, v_n(t)]^T$, in which $v_i(t) = y_i(t)/\Omega$ is the concentration of species $i$. Then the CLE is given by

$$\frac{dv(t)}{dt} = \sum_{j=1}^{m} q_j \tilde{a}_j(v) + \frac{1}{\sqrt{\Omega}} \sum_{j=1}^{m} q_j \sqrt{\tilde{a}_j(v)} \Gamma_j,$$  

where $\tilde{a}_j(v)$ are macroscopic reaction rates which can be approximated by $\tilde{a}_i(v) = \frac{1}{\Omega} a_i(\Omega v)$ in the thermodynamics limit. i.e., $\Omega \to \infty$ and $y \to \infty$ such that the concentration $v = y/\Omega$ remains constant.

### 2.1.3 Linear Noise Approximation (LNA)

The Linear Noise Approximation is another approximation to the CME, where the molecular counts are represented by continuous variables under the assumption that the system volume and the molecular counts are sufficiently large. Introduced by van Kampen [90], the LNA is derived by taking the approximation $y = \Omega v + \sqrt{\Omega} \xi$ in the CME, where $v$ is a vector of deterministic variables and $\xi$ is a vector of random variables that represents the stochastic fluctuations. Substituting this approximation in the CME and performing a Taylor series expansion about the deterministic variable $\Omega v$ and equating the terms of order $\Omega^{1/2}$ and $\Omega^0$, van Kampen shows that $v$ gives the macroscopic concentrations and the elements of $\xi$ are Gaussian random variables with the dynamics

$$\dot{v} = f(v),$$  

(2.8)
\[ \dot{\xi} = A(v)\xi + \sigma(v)\Gamma, \quad (2.9) \]

in which \( \Gamma \) is an \( m \)-dimensional white noise process,

\[
(v, t) = \sum_{i=1}^{m} q_i \tilde{a}_i(v), \\
A(v, t) = \frac{\partial f(v)}{\partial v}, \\
\sigma(v, t) = [q_1 \sqrt{\tilde{a}_1(v)}, \ldots, q_m \sqrt{\tilde{a}_m(v)}],
\]

where the function \( \tilde{a}_i(v) \) is the macroscopic reaction rate which can be approximated by \( \tilde{a}_i(v) = \frac{1}{\Omega} a_i(\Omega v) \) as \( \Omega \to \infty \) and \( y \to \infty \) such that the concentration \( v = y/\Omega \) remains constant [24].

### 2.2 Gene-regulatory networks

Gene-regulatory networks describe the interactions between genes in a cell and the proteins that regulate the expressions of these genes [2]. The regulatory proteins, known as transcription factors (TFs), can activate or repress the expression of a gene by binding to its promoter site and by either aiding or interfering with the transcription process by RNA polymerase. Through these activation and repression processes, the cell regulates the levels of proteins in the cell, which are responsible for a vast majority of cellular functions.

In the modeling point of view, a gene-regulatory network can be represented as a collection of interconnected nodes, in which each node represents a transcriptional component as shown in Figure 2-1. A transcriptional component encapsulates the process of protein production and can be viewed as a system that takes several input transcription factors and produces as output the target gene’s protein product. The process of protein production occurs in two steps:

1. Transcription, where the genetic sequence is transcribed into mRNA by RNA polymerase
2. Translation, where the mRNA is translated into proteins via ribosomes.

The regulation of proteins via TFs occurs at the transcription step where the TFs can either facilitate or interfere with the transcription process by binding to the promoter site of the gene. A TF that facilitates the production of a protein are known as an activator and a TFs that inhibits the production is known as a repressor.

Typically, it is assumed that the resources such as RNA polymerase and ribosomes are present in abundance and the protein production is represented in a one-step reaction that lumps both transcription and translation processes.

Figure 2-1: Schematic of a gene-regulatory network. Node $i$ is an input to node $j$, which is in turn an input to node $k$. 

Transcriptional Component
Chapter 3

Reduced-order model for multi-scale chemical Langevin equation with linear propensity functions

3.1 Introduction

In this chapter, we consider the problem of model order reduction of chemical Langevin equations with linear propensity functions. We introduce a reduced-order model to approximate both slow and fast variable dynamics of the system and show that the moment dynamics of the reduced-order model is in an $O(\epsilon)$-neighborhood of the moment dynamics of the full system, for moments of all orders for the slow variable and for first and second moments of the fast variable. We begin by presenting the system model and the assumptions in Section 3.2. We then introduce the reduced-order model and present the results on the error quantification in Chapter 3.3. The results in this chapter appeared in [36, 31, 34].
3.2 System model

In this chapter, we consider biomolecular systems with \( n \) species and \( m \) reactions modeled by the chemical Langevin equations of the form:

\[
\frac{dv(t)}{dt} = \sum_{j=1}^{m} q_j \tilde{a}_j(v) + \frac{1}{\sqrt{\Omega}} \sum_{j=1}^{m} q_j \sqrt{\tilde{a}_j(v)} \Gamma_j, \tag{3.1}
\]

where \( q_j \) are the stoichiometry vectors and \( \tilde{a}_j(v) \) are the macroscopic reaction vectors as described in Section 2.

Under time-scale separation conditions where the reactions take place on well-separate time-scales, we can partition the reactions into \( m_s \) slow reactions and \( m_f \) fast reactions. Then, by defining a small parameter \( \epsilon \), we can arrange the reaction rate vector in the form \( \tilde{a}(v) = [\tilde{a}_s(v), (1/\epsilon)\tilde{a}_f(v)]^T \) where \( \tilde{a}_s(v) \in \mathbb{R}^{m_s} \) represents the reaction rates of slow reactions and \( (1/\epsilon)\tilde{a}_f(v) \in \mathbb{R}^{m_f} \) represents the reaction rates of fast reactions. The corresponding stoichiometry vectors \( q_i \) can be written in the form \( q = [q_1, \ldots, q_{m_s}, q_{m_s+1}, \ldots, q_{m_s+m_f}] \) where \( q_i \) for \( i = 1, \ldots, m_s \) represent the change in the molecular counts given by the slow reactions, and \( q_i \) for \( i = m_s + 1, \ldots, m_s + m_f \) represent the change in the molecular counts given by the fast reactions. Because, chemical species often take part in both slow and fast reactions, the above separation in reaction rates does not necessarily correspond to a partitioning of the system's species into fast and slow. Often, a coordinate change is necessary to identify the slow and fast variables in the system and write it in the standard singular perturbation form [40, 13, 26]. Therefore, here we consider systems where an invertible coordinate transformation of the form \( T = [T_x^T, T_z^T]^T \) with \( T_x \in \mathbb{R}^{n_s \times n} \) and \( T_z \in \mathbb{R}^{n_f \times n} \) and the change of variables \( x = T_x v, z = T_z v \) can be used to take the system (3.1) into the singular perturbation form:

\[
\dot{x} = f_x(x, z, t) + \sigma_x(x, z, t) \Gamma_x, \quad x(0) = x_0, \tag{3.2}
\]
\[
\epsilon \dot{z} = f_z(x, z, t, \epsilon) + \sqrt{\epsilon} \sigma_z(x, z, t, \epsilon) \Gamma_z, \quad z(0) = z_0, \tag{3.3}
\]

where \( x \in D_x \subseteq \mathbb{R}^{n_s} \) is the slow variable and \( z \in D_z \subseteq \mathbb{R}^{n_f} \) is the fast variable. \( \Gamma_z \) is
a $d_x$-dimensional white noise process. Let $\Gamma_f$ be a $d_f$-dimensional white noise process, while $\Gamma_z$ is a $(d_x + d_f)$-dimensional white noise process.

We refer to the system (3.2)–(3.3) as the full system and assume that the system satisfies the following assumptions.

**Assumption 3.1.** The functions $f_x(x, z, t)$ and $f_z(x, z, t, \epsilon)$ are affine functions of the state variables $x$ and $z$, i.e., we can write $f_x(x, z, t) = A_1 x + A_2 z + A_3(t)$, where $A_1 \in \mathbb{R}^{n_x \times n_x}$, $A_2 \in \mathbb{R}^{n_x \times n_f}$ and $A_3(t) \in \mathbb{R}^{n_x}$, $f_z(x, z, t, \epsilon) = B_1 x + B_2 z + B_3(t) + \alpha(\epsilon)(B_4 x + B_5 z + B_6(t))$, where $B_1, B_4 \in \mathbb{R}^{m \times n_x}$, $B_2, B_5 \in \mathbb{R}^{n_f \times n_x}$, $B_3(t), B_6(t) \in \mathbb{R}^{n_f}$, $A_3(t)$ and $B_3(t)$ are continuously differentiable functions, and $\alpha(\epsilon)$ is a continuously differentiable function with $\alpha(0) = 0$.

**Assumption 3.2.** Let $\Phi(x, z, t) = \sigma_x(x, z, t)\sigma_x(x, z, t)^T$, $\Lambda(x, z, t, \epsilon) = \sigma_x(x, z, t, \epsilon)\sigma_x(x, z, t, \epsilon)^T$, and $\Theta(x, z, t, \epsilon) = \sigma_x(x, z, t, \epsilon)[\sigma_x(x, z, t) 0 ]^T$. Then, we assume that $\Phi(x, z, t)$, $\Lambda(x, z, t, \epsilon)$, and $\Theta(x, z, t, \epsilon)$ are affine functions of $x$ and $z$, and that $\lim_{\epsilon \to 0} \Lambda(x, z, t, \epsilon) < \infty$ and $\lim_{\epsilon \to 0} \Theta(x, z, t, \epsilon) < \infty$ for all $x$, $z$ and $t$. Furthermore, we assume that the functions $\Phi(x, z, t)$, $\Lambda(x, z, t, \epsilon)$, and $\Theta(x, z, t, \epsilon)$ are continuously differentiable in $t$ and $\epsilon$.

**Assumption 3.3.** Matrix $B_2$ is Hurwitz.

We further assume that the system (3.2)–(3.3) admits a unique well-defined solution on a finite time interval. Sufficient conditions for the existence and uniqueness of solutions of stochastic differential equations are given by the Lipschitz continuity and bounded growth of system functions [64]. However, the class of systems considered in this work includes systems of the form where the diffusion term is a square-root function of the state variables, as Assumption 3.2 requires the squared diffusion terms to be linear functions of the state variables. Therefore, such systems may not guarantee the Lipschitz continuity conditions for the diffusion coefficient. For this type of systems, a set of sufficient conditions that guarantee the existence of solutions can be found in [16]. We also note that the existence of a well-defined solution for general chemical Langevin equations is an on-going area of research [91, 74].
We note the class of biomolecular systems that satisfy Assumptions 3.1 and 3.2 are systems where the propensity functions $\tilde{a}(y)$ are linear, i.e., unimolecular reactions.

In the next section, we introduce the reduced-order system and present the results on the error quantification between the full and reduced-order systems.

3.3 Results

3.3.1 Reduced-order model

We introduce a reduced-order model by setting $\epsilon = 0$ in the full system (3.2)–(3.3), as in the case of deterministic singular perturbation theory. Under Assumption 3.2, $\epsilon = 0$ leads to the algebraic equation $f_z(x, z, t, 0) = B_1 x + B_2 z + B_3(t) = 0$, for which, Assumption 3.3 guarantees the existence of a unique global solution $z = \gamma(x, t)$, given by

$$\gamma(x, t) = -B_2^{-1}(B_1 x + B_3(t)). \quad (3.4)$$

Upon substitution of $z = \gamma(x, t)$ into (3.2), we obtain the reduced slow system

$$\dot{x} = f_x(\bar{x}, \gamma(\bar{x}, t), t) + \sigma_x(\bar{x}, \gamma(\bar{x}, t), t) \Gamma_x, \quad \bar{x}(0) = x_0, \quad (3.5)$$

which only depends on $\bar{x}$.

We assume that system (3.5) has a unique well-defined solution on a finite time interval $[0, t_1]$.

Next, we define a candidate approximation for the fast variable dynamics in the form

$$\bar{z}(t) = \gamma(\bar{x}(t), t) + g(\bar{x}(t), t)N, \quad (3.6)$$

where $N \in \mathbb{R}^d$ is a random vector whose components are independent standard normal
random variables, and \( g(\tilde{x}(t), t) : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_d} \) is a function that satisfies the Lyapunov equation

\[
g(\tilde{x}(t), t)g(\tilde{x}(t), t)^T B_2^T + B_2 g(\tilde{x}(t), t)g(\tilde{x}(t), t)^T = -\Lambda(\tilde{x}, \gamma(\tilde{x}(t), t), t, 0). \tag{3.7}
\]

We call equation (3.6) the reduced fast system.

### 3.3.2 Preliminary results

In this section, we derive some preliminary results on the moment dynamics that will lead to the error quantification between the full and reduced-order systems. To this end, we first introduce the notation used to denote the moment dynamics (notation adapted from [78, 77]). Consider the vectors \( x = [x_1, \ldots, x_{n_x}]^T \) and \( k = (k_1, \ldots, k_{n_x}) \) where \( x_i, k_i \in \mathbb{R} \) for \( i = 1, \ldots, n_x \). Let \( x^{(k)} = x_1^{k_1} x_2^{k_2} \ldots x_{n_x}^{k_{n_x}} \). Then \( \mathbb{E}[x^{(k)}] \) denotes the moment of \( x \) corresponding to the vector \( k \), where the order of the moment is \( \sum_{i=1}^{n_x} k_i \).

Next, in order to represent a set of moments of order \( P \), we define the following sets \( \mathcal{K}_r^P = \{(k_1, \ldots, k_r) \in \mathbb{Z}_{\geq 0}^r | \sum_{i=1}^r k_i = P \}, \mathcal{G}_r^P = \{(c_1, \ldots, c_r) \in \mathbb{Z}_r^r | \sum_{i=1}^r c_i \leq P \}. \) Then, considering the full system in (3.2)-(3.3), denote the state vectors by \( x = [x_1, \ldots, x_{n_x}]^T \) and \( z = [z_1, \ldots, z_{n_f}]^T \). We then have the following claim.

**Claim 3.1.** Under Assumption 3.1 - 3.2, the moment dynamics of the full system in (3.2)-(3.3) can be written in the singular perturbation form:

\[
d\mathbb{E}[x^{(k)}] = \sum_{i \in \mathcal{G}_r^P} C_{1i}(t)\mathbb{E}[x^{(i)}] + \sum_{i \in \mathcal{G}_r^P} \sum_{j \in \mathcal{G}_r^P} C_{2ij}(t)\mathbb{E}[z^{(j)}] x^{(i)}\mathbb{E}[x^{(j)}], \forall k \in \mathcal{K}_r^P \tag{3.8}
\]

\[
ed\mathbb{E}[z^{(g)}] = \sum_{a \in \mathcal{G}_r^P} D_{1a}(t, \epsilon)\mathbb{E}[z^{(a)}] + \sum_{a \in \mathcal{G}_r^P} \sum_{b \in \mathcal{G}_r^P} D_{2ab}(t, \epsilon)\mathbb{E}[x^{(a)}] z^{(b)}\mathbb{E}[z^{(b)}], \forall g \in \mathcal{K}_r^P \tag{3.9}
\]

\[
ed\mathbb{E}[z^{(k)} x^{(k)}] = \sum_{u \in \mathcal{G}_r^P} F_{1u}(t, \epsilon)\mathbb{E}[x^{(u)}] + \sum_{a \in \mathcal{G}_r^P} F_{2a}(t, \epsilon)\mathbb{E}[z^{(a)}] + \sum_{q=2}^P \sum_{r=1}^q \sum_{k \in \mathcal{G}_r^P} \sum_{s \in \mathcal{G}_r^P} F_{3qrs}(t, \epsilon)\mathbb{E}[z^{(k)} x^{(s)}], \tag{3.10}
\]
for $k_x \in \mathcal{K}_{n_x}^{Q_x}$ and $k_z \in \mathcal{K}_{n_z}^{Q_z}$, where $Q_x + Q_z = P$, and for appropriate continuous functions $C_{11}(t), C_{2jl}(t)$ and continuously differentiable functions $D_{1a}(t), D_{2bc}(t), F_{2a}(t, \epsilon), F_{11}(t, \epsilon), F_{3qrks}(t, \epsilon)$, for $i \in \mathcal{G}_{n_x}^P$, $l \in \mathcal{G}_{n_f}^1$ and $j \in \mathcal{G}_{n_z}^{P-1}$, $a \in \mathcal{G}_{n_f}^P$, $c \in \mathcal{G}_{n_z}^1$, $b \in \mathcal{G}_{n_f}^{P-1}$, $u \in \mathcal{G}_{n_z}^P$, $q = 2, \ldots, P$, $r = 1, \ldots, q$, $k \in \mathcal{G}_{n_f}^r$, $s \in \mathcal{G}_{n_z}^{r-1}$, $P = \{1, \ldots, N\}$ where $N \in \mathbb{Z}_{\geq 0}$.

**Proof.** See Appendix A.1.\qed

We note that the proof of this claim is based on the linearity conditions in Assumption 3.1 and 3.2, which guarantee that the set of resulting moment equations will be closed.

Next, we derive the moment dynamics of the reduced-order system (3.5)-(3.6). For this, denote the state vector of the reduced slow system by $\bar{x} = [x_1, \ldots, x_{n_s}]^T$ and let $\gamma(\bar{x}, t) = [\gamma_1(\bar{x}, t), \ldots, \gamma_{n_f}(\bar{x}, t)]$, $g(\bar{x}, t) = [g_1(\bar{x}, t), \ldots, g_{Pz+Pf}(\bar{x}, t)]$ where $g_{*i}(\bar{x}, t)$ denotes the $i$th column of $g(\bar{x}, t)$. Then, we have the following claim.

**Claim 3.2.** Under Assumptions 3.1 - 3.2, the moment dynamics of the reduced slow system in (3.5) are given by

$$
\frac{d\mathbb{E}[\bar{x}^{(k)}]}{dt} = \sum_{i \in \mathcal{G}_{n_s}^P} C_{11}(t)\mathbb{E}[\bar{x}^{(i)}] + \sum_{l \in \mathcal{G}_{n_f}^1} \sum_{j \in \mathcal{G}_{n_z}^{P-1}} C_{2jl}(t)\mathbb{E}[\gamma(\bar{x}, t)^{(l)}\bar{x}^{(j)}], \quad \forall k \in \mathcal{K}_{n_x}^{Pz},
$$

and the dynamics of the first and second moments of the reduced fast system in (3.6) are given by

$$
\mathbb{E}[^{(h)}] = \mathbb{E}[\gamma(\bar{x}, t)^{(h)}] + (P_z - 1)\mathbb{E}[\sum_{l=1}^{dz+dz} g_{*l}(\bar{x}, t)^{(h)}], \quad \forall h \in \mathcal{K}_{n_f}^{Pz},
$$

where the functions $C_{11}(t) : \mathbb{R} \to \mathbb{R}$ and $C_{2jl}(t) : \mathbb{R} \to \mathbb{R}$ satisfies equation (3.8) in Claim 3.1, $P_z = \{1, \ldots, N\}$ where $N \in \mathbb{Z}_{\geq 0}$ and $P_z = \{1, 2\}$.

**Proof.** See Appendix A.2.\qed

36
Similar to Claim 3.1, the derivation of the moment dynamics in Claim 3.2 utilizes linearity conditions in Assumption 3.1 and 3.2 to obtain a set of closed moment equations.

Next, we analyze the set of moment equations obtained by setting $\epsilon = 0$ in the moment dynamics of the full system given in Claim 3.1.

**Claim 3.3.** Setting $\epsilon = 0$ in the moment dynamics of the full system in (3.8)-(3.10) yields the following reduced-order system for all moments of the slow variable and up to second order moments of the fast variable:

$$\frac{d\mathbb{E}[x^{(k)}]}{dt} = \sum_{i \in \mathbb{P}_x} C_{1i}(t)\mathbb{E}[x^{(i)}] + \sum_{l \in \mathbb{P}_y} \sum_{j \in \mathbb{P}_y^{-1}} C_{2lj}(t)\mathbb{E}[\gamma(x, t)^{(l)}x^{(j)}], \quad \forall k \in \mathbb{K}_{n_x}^{P_x},$$

(3.13)

$$\mathbb{E}[x^{(h)}] = \mathbb{E}[\gamma(x, t)^{(h)}] + (P_x - 1)\mathbb{E}\left[\sum_{l=1}^{d_x + d_f} g_{hl}(\bar{x}, t)^{(h)}\right], \quad \forall h \in \mathbb{K}_{n_f}^{P_x},$$

(3.14)

where the functions $C_{1i}(t) : \mathbb{R} \to \mathbb{R}$ and $C_{2lj}(t) : \mathbb{R} \to \mathbb{R}$ satisfies equation (3.8) in Claim 3.1, $P_x = \{1, \ldots, N\}$ where $N \in \mathbb{Z}_{>0}$ and $P_z = \{1, 2\}$.

**Proof.** See Appendix A.3.

In the next section, we use the moment dynamics derived in this section to prove that the error between moments of the full system and the reduced system are of $O(\epsilon)$.

### 3.3.3 Main results

**Lemma 3.1.** Consider the full system in (3.2)-(3.3), the reduced system in (3.5)-(3.6), the moment dynamics of the full system in (3.8)-(3.10) and the moment dynamics of the reduced system in (3.11). We have that, under Assumptions 3.1 - 3.3, the commutative diagram in Figure 3-1 holds.

**Proof.** Proof follows from Claim 3.1, Claim 3.2 and Claim 3.3.
The proof of this Theorem (presented in Appendix A.4) is based on applying the Tikhonov’s theorem to the moment dynamics of the full and the reduced systems. From Claim 3.1 - 3.3 in Section 3.3.2 and Lemma 3.1, it can be seen that the moment dynamics of the full system are in the standard singular perturbation form, and that setting $\epsilon = 0$ in the full moment dynamics yields the moment dynamics of the reduced-order system. This holds for moments of all orders for the slow variables and up to second order moments for the fast variables. As the moment dynamics are deterministic, we then apply the Tikhonov’s theorem to demonstrate the convergence of the moments of the reduced-order system to the moments of the full system, as $\epsilon$ decreases. The stability conditions of the slow manifold of the full moment dynamics required for the application of the Tikhonov’s theorem are guaranteed by Assumption 3.3.

From the reduced-order approximations given in equations (3.5)–(3.6), we note the similarity with the reduced-order model obtained by singular perturbation theory for deterministic systems [45]. In particular, the slow variable’s dynamics are well
approximated by substituting the expression of the slow manifold given by \( z = \gamma(x, t) \) in equation (3.4) into the slow variable’s dynamics given in equation (3.2). This implies that for this class of systems, the slow variable approximation can be obtained in the same manner as in the deterministic singular perturbation method.

By contrast, from expression (3.6) we note that the fast variable approximation contains the term \( g(\bar{x}, t)N \), which is in addition to the slow manifold expression \( \gamma(\bar{x}, t) \) that would be obtained with direct application of deterministic singular perturbation theory. This additional term is required in order to account for the noise of the fast variables. In fact, considering the system in the fast time-scale \( \tau = t/\epsilon \), we see that the SDE of the fast variable is given by

\[
\frac{dz}{d\tau} = f_z(x, z, t, \epsilon) + \sigma_z(x, z, t, \epsilon)\bar{\Gamma}_z, \tag{3.17}
\]

where \( \bar{\Gamma}_z \) represents \( \Gamma_z \) in the fast-time scale, i.e., \( \bar{\Gamma}_z(\tau) = \sqrt{\epsilon} \Gamma_z(t) \) as shown in [93, p.173]. For the case where the diffusion term is of the order \( \sqrt{\epsilon} \), the term \( \sigma_z(x, z, t, \epsilon) \) is independent of \( \epsilon \) and thus \( \sigma_z(x, z, t, 0) \neq 0 \). This shows that the fast variable is subject to noise, given by the diffusion term \( \sigma_z(x, z, t, \epsilon) \), and thus the expression \( \gamma(x, t) \) does not provide an adequate approximation for the noise on \( z \).

The noise in the fast variable can be “neglected” in the slow variable approximation because the slow subsystem “filters out” the noise from the fast variable. Such noise must instead be considered to approximate the noise properties of the fast variable, as we illustrate in the following example. Consider the system

\[
\dot{x} = -a_1 x + a_2 z + v_1 \Gamma_1, \tag{3.18}
\]
\[
\epsilon \dot{z} = -z + v_2 \sqrt{\epsilon} \Gamma_2, \tag{3.19}
\]

where \( a_1, a_2 > 0 \).

Setting \( \epsilon = 0 \), we obtain the system:

\[
\dot{x} = -a_1 \bar{x} + v_1 \Gamma_1, \tag{3.20}
\]
\[ z = \gamma(\bar{x}, t) = 0. \] (3.21)

To analyze the error of this approximation, we can directly calculate the steady state moments for both the full and reduced-order systems using their linearity. This yields

\[
\begin{align*}
\mathbb{E}[x^2] &= \frac{a_2^2\nu_2}{2a_1} \frac{\epsilon}{1 + a_1\epsilon} + \frac{\nu_1^2}{2a_1}, \\
\mathbb{E}[z^2] &= \frac{\nu_2^2}{2}, \\
\mathbb{E}[\bar{x}^2] &= \frac{\nu_1^2}{2a_1},
\end{align*}
\]

It is seen that \( \mathbb{E}[x^2] \) converges to \( \mathbb{E}[\bar{x}^2] \) as \( \epsilon \) approaches zero, however, \( \mathbb{E}[z^2] \) remains constant as \( \epsilon \) goes to zero. That is, the reduced-order system (3.20)-(3.21) obtained by setting \( \epsilon = 0 \) provides a good approximation for the slow variable in terms of the second moment, but it is not a good approximation for the fast variable dynamics. This is due to the fact that the \( x \)-subsystem is not affected by the noise \( \Gamma_2 \) as \( \epsilon \) tends to zero, which can be explained by considering the power spectra and frequency response of the \( x \) and \( z \) subsystems.

Using the frequency response from input \( \Gamma_2 \) to the output \( z \) of the \( z \)-subsystem, given by \( H_{z1}(j\omega) = \frac{1}{j\omega + 1/\epsilon} \) we can calculate the power spectrum of \( z \) as \( S_{zz}(\omega) = \frac{(\nu_2/\sqrt{\epsilon})^2}{\omega^2 + (1/\epsilon)^2} \), which is illustrated in Figure 3-2. It can be seen that as \( \epsilon \) approaches zero, the magnitude of \( S_{zz}(\omega) \) decreases at low frequencies but increases at high frequencies, in a way that the variance of \( z \) remains constant. However, considering the frequency response from \( z \) to \( x \) of the \( x \)-subsystem, given by \( H_{xz}(j\omega) = \frac{a_2}{j\omega + a_1} \), we see that the \( x \)-subsystem is a low-pass filter with a cut-off frequency of \( a_1 \) that is independent of \( \epsilon \) (Figure 3-2). Therefore, \( x \) only selects the low frequency components of signal \( z \), which decrease with \( \epsilon \), leading to a decrease in the variance of signal \( x \) as \( \epsilon \) decreases. Thus, the reduced-order system obtained by setting \( \epsilon = 0 \) provides a good approximation for the slow variable dynamics. However, as the variance of \( z \) remains constant as \( \epsilon \) decreases, the expression \( \bar{z} = \gamma(\bar{x}, t) \) by itself does not provide a good approximation for the fast variable stochastic dynamics.

The stochastic differential equations considered in this chapter can also arise in
areas such as finance, in addition to biomolecular system models. Therefore, next we illustrate the application of the model reduction approach on an academic example.

### 3.4 Academic example

We consider the following system, which takes a similar form to the SDEs that appear in affine term structure models in finance [16]:

\[
\begin{align*}
\dot{x} &= -2x + z + 10 + \sqrt{2z + 1} \Gamma_1, \\
\epsilon \dot{z} &= -z + 15 + \sqrt{\epsilon(2z + 1)} \Gamma_2.
\end{align*}
\]  

This system satisfies the Assumptions 3.1 - 3.3, and using the results of [16] it can be verified that there exists a unique, well-defined solution. Setting \(\epsilon = 0\), we obtain the slow manifold \(z = 15\). This yields the following reduced-order model for the slow
variable:

\[ d\bar{x} = -2\bar{x} + 25 + \sqrt{31}\Gamma_1. \] \hfill (3.24)

Based on (3.6), the fast variable approximation for this system is of the form

\[ z = 15 + g(x)N, \] \hfill (3.25)

where \( g(x)g(x)^T(-1) + (-1)g(x)g(x)^T = 31 \) and \( N \) is a standard normal random variable. After solving for \( g(x) \), the fast variable approximation is given by

\[ \bar{z} = 15 + \sqrt{15.5}N. \] \hfill (3.26)

Simulations of the full and the reduced-order systems were performed using the Euler-Maruyama method \cite{37} for stochastic differential equations and the sample means were calculated using 500,000 realizations. Figure 3-3 illustrates the second and third order moments of the slow variable and second order moments of the fast variable for the full and reduced-order systems. It can be seen that as \( \epsilon \) decreases the moments of the full system tends to the moments of the reduced-order system.

![Figure 3-3: Moments of the full and reduced systems. (a) Second moments of the slow variable. (b) Third moments of the slow variable. (c) Second moments of the fast variable.](image)

By virtue of Theorem 3.1, the reduced-order model (3.24) provides a good approximation of the higher moments for the slow variable, as illustrated in Figure 3-3.
However, for the fast variable, only the first and second moments are well approximated, and there is no guarantee that the higher order moments are also approximated well, as we show by analyzing the third order moments of the system considered in this section.

To calculate the third order moments, we represent the fast variable dynamics of the above system (3.22)–(3.23) in the form

\[ \epsilon dz = c_1 z + c_2 + \sqrt{\epsilon(d_1 z + d_2)} \Gamma_2. \]

Then, the third order moment dynamics are written as

\[
\begin{align*}
\epsilon \frac{d\mathbb{E}[z]}{dt} &= c_1 \mathbb{E}[z] + c_2, \\
\epsilon \frac{d\mathbb{E}[z^2]}{dt} &= 2c_1 \mathbb{E}[z^2] + 2c_2 \mathbb{E}[z] + d_1 \mathbb{E}[z] + d_2, \\
\epsilon \frac{d\mathbb{E}[z^3]}{dt} &= 3c_1 \mathbb{E}[z^3] + 3c_2 \mathbb{E}[z^2] + 3d_1 \mathbb{E}[z^2] + 3d_2 \mathbb{E}[z].
\end{align*}
\]

Setting \( \epsilon = 0 \), we obtain

\[
\begin{align*}
\mathbb{E}[z] &= \frac{-c_2}{c_1}, \\
\mathbb{E}[z^2] &= \frac{2c_2^2 + d_1 c_2 - d_2 c_1}{2c_1^2}, \\
\mathbb{E}[z^3] &= \frac{3c_1 c_2 d_2 + c_1 d_1 d_2 - 2c_2^3 - 3c_2^2 d_1 - c_2 d_1^2}{2c_1^3}.
\end{align*}
\]

The reduced fast system is given by \( \ddot{z} = \gamma(\bar{x}, t) + g(\bar{x}, t) N(0, 1) \), where \( \gamma(\bar{x}, t) = \frac{-c_2}{c_1} \) and \( g(\bar{x}, t) = \frac{d_1 \gamma(\bar{x}, t) + d_2}{-2c_1} \). Calculating the moment dynamics for the reduced fast system we obtain

\[
\begin{align*}
\mathbb{E}[\ddot{z}] &= \gamma(\mathbb{E}[\bar{x}], t) = \frac{-c_2}{c_1}, \\
\mathbb{E}[\dot{z}^2] &= \mathbb{E}[\gamma(\bar{x}, t)^2] + \mathbb{E}[g(\bar{x}, t)^2] = \frac{2c_2^2 + d_1 c_2 - d_2 c_1}{2c_1^2}, \\
\mathbb{E}[\dot{z}^3] &= \mathbb{E}[\gamma(\bar{x}, t)^3] + 3 \mathbb{E}[\gamma(\bar{x}, t)g(\bar{x}, t)^2] = \frac{c_2(3c_1 d_2 - 2c_2^2 - 3c_2 d_1)}{2c_1^3}.
\end{align*}
\]
Considering the equations for the slow manifold in (3.27) - (3.29) and the moments of the reduced fast system (3.30)-(3.32), we have that \( \|\mathbb{E}[z] - \mathbb{E}[\tilde{z}]\| = 0, \|\mathbb{E}[z^2] - \mathbb{E}[\tilde{z}^2]\| = 0 \), however, \( \|\mathbb{E}[z^3] - \mathbb{E}[\tilde{z}^3]\| = \frac{d_1(c_1d_2-c_2d_1)}{2c_1^2} \), which is different from zero. Therefore, it follows that setting \( \epsilon = 0 \) in the third moments of the fast variable does not yield the third moment of the reduced fast system.

From the general form of the moments in (3.30)-(3.32) it follows that the terms \( \gamma(x, t) \) and \( g(x, t) \) are not sufficient to approximate the third moment. This suggests that approximation of higher order moments of the fast variable would require additional terms in the reduced fast system. However, in many applications, particularly biomolecular systems, the common measures of noise are coefficient of variation and signal-to-noise ratio, which are functions of only the mean and the variance. Thus, the first and second moments provide sufficient information for analysis of these systems.

### 3.5 Conclusion

In this chapter, we considered the problem of model order reduction for a chemical Langevin equations in singular perturbation form. We introduced a reduced-order model that approximates both the slow and fast dynamics of the full system and can be obtained by solving a set of algebraic equations. For the slow variable approximation, it was shown that the error between the moments of the reduced system and moments of the full system are of \( O(\epsilon) \). For the fast variable approximation, it was shown that the first and the second moments of the reduced system are within an \( O(\epsilon) \)-neighborhood of the first and second moments of the full system, respectively.

The form of the stochastic differential equations considered in this chapter can also arise in other fields such as finance. We therefore provided an illustration of the application of our results with an example that takes a similar form to the systems seen in finance. For this example, we derived the reduced-order model and verified the results of error convergence through numerical simulations.
Chapter 4

Reduced-order model for multi-scale
Linear Noise Approximation: The
stochastic tQSSA$^+$

4.1 Introduction

In this chapter, we consider the model order reduction of Linear Noise Approximation. Similar to the previous chapter, we obtain a reduced-order LNA that approximates both the slow and fast variables in the system, under time-scale separation conditions. These results, in particular, provide a rigorous justification to the accuracy of LNA models derived using the stochastic total quasi-steady state approximation (tQSSA). Since, in contrast to the stochastic tQSSA, our reduced-order model also provide approximations for the fast variable stochastic properties, we term our method the stochastic tQSSA$^+$. We begin by introducing the system model and the assumptions considered. Next, we present the reduced-order model and the results on the error quantification between the original and reduced-order moment dynamics. The results presented in this chapter appeared in [33, 35].

45
4.2 System model

In this chapter, we consider biomolecular systems with \( m \) reactions and \( n \) species whose dynamics are modeled by the Linear Noise Approximation of the form

\[
\dot{v} = f(v), \\
\dot{\xi} = A(v)\xi + \sigma(v)\Gamma,
\]

in which \( \Gamma \) is an \( m \)-dimensional white noise process, \( f(v) = \sum_{i=1}^{m} q_i \tilde{a}_i(v), A(v) = \frac{\partial f(v)}{\partial v} \)
and \( \sigma(v) = [q_1 \sqrt{\tilde{a}_1(v)}, \ldots, q_m \sqrt{\tilde{a}_m(v)}] \) and the function \( \tilde{a}_i(v) \) is the macroscopic reaction rate, as described in Section 2.

Similar to Chapter 4, considering the case of time-scale separation where the chemical reactions take place on two well-separated time-scales, we can separate the \( m \) reactions into slow and fast groups. Let \( m_s \) be the number of slow reactions and \( m_f \) be the number of fast reactions where \( m_s + m_f = m \). Next, by defining a small parameter \( \epsilon \), we can arrange the reaction rate vector in the form \( \tilde{a}(v) = [\tilde{a}_s(v), (1/\epsilon)\tilde{a}_f(v)]^T \) where \( \tilde{a}_s(v) \in \mathbb{R}^{m_s} \) represents the reaction rates of slow reactions and \( (1/\epsilon)\tilde{a}_f(v) \in \mathbb{R}^{m_f} \) represents the reaction rates of fast reactions. The corresponding stoichiometry vectors \( q_i \) can be written in the form \( q = [q_1, \ldots, q_{m_s}, q_{m_s+1}, \ldots, q_{m_s+m_f}] \) where \( q_i \) for \( i = 1, \ldots, m_s \) represent the change in the molecular counts given by the slow reactions, and \( q_i \) for \( i = m_s + 1, \ldots, m_s + m_f \) represent the change in the molecular counts given by the fast reactions. As chemical species often take part in both slow and fast reactions, we often require a coordinate change bring the system in to the standard singular perturbation form \([40, 13, 26]\). Therefore, here we consider systems in which the species can be partitioned into \( n_s \) slow variables and \( n_f \) fast variables with \( n_s + n_f = n \), according to the following claim:

Claim 4.1. Assume that there is an invertible matrix \( T = [T_x^T, T_z^T]^T \) with \( T_x \in \mathbb{R}^{n_s \times n} \) and \( T_z \in \mathbb{R}^{n_f \times n} \) such that the change of variables \( x = T_xv, z = T_zv \), takes the system
(4.1) into the singular perturbation form

\[ \dot{x} = f_x(x, z, t), \quad (4.3) \]
\[ \epsilon \dot{z} = f_z(x, z, t, \epsilon). \quad (4.4) \]

Then, the change of variables \( \psi_z = T_{x}\xi, \psi_z = T_{x}\xi \) transforms system (4.2), into the singular perturbation form

\[ \dot{\psi}_x = S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z + \sigma_x(x, z, t)\Gamma_x, \quad (4.5) \]
\[ \epsilon \dot{\psi}_z = F_x(x, z, t, \epsilon)\psi_x + F_z(x, z, t, \epsilon)\psi_z + \sigma_z(x, z, t, \epsilon)\Gamma_z, \quad (4.6) \]

where \( \Gamma_x \) is an \( m_x \)-dimensional white noise process, \( \Gamma_z = [\Gamma_x^T, \Gamma_f^T]^T \), where \( \Gamma_f \) is an \( m_f \)-dimensional white noise process and

\[
S_x(x, z, t) = \frac{\partial f_x(x, z, t)}{\partial x}, \quad S_z(x, z, t) = \frac{\partial f_z(x, z, t)}{\partial z}, \\
F_x(x, z, t, \epsilon) = \frac{\partial f_x(x, z, t, \epsilon)}{\partial x}, \quad F_z(x, z, t, \epsilon) = \frac{\partial f_z(x, z, t, \epsilon)}{\partial z}, \\
\sigma_x(x, z, t) = T_x\left[ q_1 \sqrt{a_{s1}(T^{-1}[x^T, z^T]^T, t), \ldots, q_{m_x} \sqrt{a_{sm_x}(T^{-1}[x^T, z^T]^T, t)} \right], \\
\sigma_z(x, z, t, \epsilon) = T_x\left[ \epsilon \left[ q_1 \sqrt{a_{s1}(T^{-1}[x^T, z^T]^T, t), \ldots, q_{m_x} \sqrt{a_{sm_x}(T^{-1}[x^T, z^T]^T, t)} \right]^T_T z \right. \\
\left. T_{x+1} \sqrt{\epsilon a_{f1}(T^{-1}[x^T, z^T]^T, \epsilon), \ldots, q_{m_x+m_f} \sqrt{\epsilon a_{fm_f}(T^{-1}[x^T, z^T]^T, \epsilon)} \right]^T_T z \right].
\]

**Proof.** See Appendix B.1. \( \square \)

Following the results of Claim 4.1, here, we consider biomolecular systems where the Linear Noise Approximation model can be written in the standard singular perturbation form

\[ \dot{x} = f_x(x, z, t), \quad x(0) = x_0, \quad (4.7) \]
\[ \epsilon \dot{z} = f_z(x, z, t, \epsilon), \quad z(0) = z_0, \quad (4.8) \]
\[ \dot{\psi}_x = S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z + \sigma_x(x, z, t)\Gamma_x, \quad \psi_x(0) = \psi_{x0}, \quad (4.9) \]

47
\[
\epsilon \dot{\psi}_z = F_x(x, z, t, \epsilon) \psi_x + F_z(x, z, t, \epsilon) \psi_z + \sigma_x(x, z, t, \epsilon) \Gamma_z, \quad \psi_z(0) = \psi_{z0}, \tag{4.10}
\]

where \( x \in \mathbb{R}^{n_x}, \psi_x \in \mathbb{R}^{n_x} \) are the slow variables and \( z \in \mathbb{R}^{n_f}, \psi_z \in \mathbb{R}^{n_f} \) are the fast variables. \( \Gamma_x \) is an \( m_x \)-dimensional white noise process. Then, \( \Gamma_z = [\Gamma_z^T, \Gamma_f^T]^T \), where \( \Gamma_f \) is an \( m_f \)-dimensional white noise process.

We refer to the system (4.7) - (4.10) as the \textit{full system}. We make the following assumptions for \( x \in \mathbb{R}^{n_x} \) and \( z \in \mathbb{R}^{n_f} \):

**Assumption 4.1.** The functions \( f_x(x, z, t) \), \( f_z(x, z, t, \epsilon) \) are twice continuously differentiable. The Jacobian \( \frac{\partial f_x(x, z, t, \epsilon)}{\partial x} \) has continuous first and second partial derivatives with respect to its arguments.

**Assumption 4.2.** The matrix-valued functions \( \sigma_x(x, z, t) \sigma_x(x, z, t)^T \), \( \sigma_z(x, z, t, \epsilon) \sigma_z(x, z, t, \epsilon)^T \) are continuously differentiable. Furthermore, let \( \Lambda(x, z, t, \epsilon) = \frac{\sigma_z(x, z, t, \epsilon) \sigma_z(x, z, t, \epsilon)^T}{\epsilon} \). We have that \( \sigma_z(x, z, t, 0) = 0 \) and \( \Lambda(x, z, t, 0) \) is bounded for given \( x, z, t \).

**Assumption 4.3.** There exists an isolated real root \( z = \gamma_1(x, t) \), for the equation \( f_z(x, z, t, 0) = 0 \), for which, the matrix \( \frac{\partial f_z(x, z, t, 0)}{\partial z} \big|_{z=\gamma_1(x,t)} \) is Hurwitz (i.e. eigenvalues of the matrix have strictly negative real parts), uniformly in \( x \) and \( t \). Furthermore, we have that the first partial derivative of \( \gamma_1(x, t) \) is continuous with respect to its arguments. Also, the initial condition \( z_0 \) is in the region of attraction of the equilibrium point \( z = \gamma_1(x_0, 0) \) for the system \( \frac{dz}{dr} = f_z(x_0, z, 0, 0) \).

**Assumption 4.4.** The system \( \dot{x} = f_x(x, \gamma_1(x, t), t) \) has a unique bounded solution for \( t \in [0, t_1] \).

In the next section, we present the reduced-order model that we define to approximate the slow and fast variables when \( \epsilon \) is small in the full system (4.7) - (4.10), and quantify the error between the moment dynamics of the full and the reduced systems.
4.3 Results

4.3.1 Reduced-order model

To define the reduced system we follow a similar approach to the singular perturbation theory [45] by setting $\epsilon = 0$ in the full system (4.7) - (4.10). This yields

\[ f_z(x, z, t, 0) = 0, \]
\[ F_z(x, z, t, 0) \dot{\psi}_x + F_z(x, z, t, 0) \psi_z = 0. \]

Let $z = \gamma_1(x, t)$ be an isolated root of equation (4.11). Then, it follows that the unique solution to equation (4.12) is

\[ \psi_z = -F_z(x, \gamma_1(x, t), t, 0)^{-1} F_z(x, \gamma_1(x, t), t, 0) \psi_z. \]

Note that the invertibility of matrix $F_z$ follows from Assumption 4.3 where it is assumed that $F_z(x, \gamma_1(x, t), t, 0) = \frac{\partial f_z(x, z, t, 0)}{\partial z} |_{z=\gamma_1(x, t)}$ is Hurwitz. Let $\gamma_2(x, t) = -F_z(x, \gamma_1(x, t), t, 0)^{-1} F_z(x, \gamma_1(x, t), t, 0)$. Substituting $z = \gamma_1(x, t)$ and $\psi_z = \gamma_2(x, t) \psi_z$ in equations (4.7) and (4.9), we obtain the following candidate approximation for the slow variable dynamics:

\[ \dot{x} = f_z(x, \gamma_1(x, t), t), \quad x(0) = x_0, \]
\[ \dot{\psi}_z = S(x, t) \psi_z + \sigma_z(\bar{x}, \gamma_1(x, t), t) \Gamma_z, \quad \psi_z(0) = \psi_{z0}. \]

where $S(x, t) = S_x(x, \gamma_1(x, t), t) + S_z(x, \gamma_1(x, t), t) \gamma_2(x, t)$.

Next, to introduce a candidate approximation for the fast variable dynamics we define:

\[ \bar{z} = \gamma_1(x, t), \]
\[ \bar{\psi}_z = \gamma_2(x, t) \bar{\psi}_z + g(\bar{x}, t) N(0, 1), \]

where $N(0, 1) \in \mathbb{R}^d$ is a vector of standard normal random variables and $g(x, t)$ :
\[ F_{\varepsilon}(x, \gamma_1(x, t), t, 0)g(x, t)g(x, t)^T + g(x, t)g(x, t)^TF_{\varepsilon}(x, \gamma_1(x, t), t, 0)^T = -\Lambda(x, \gamma_1(x, t), t, 0), \] (4.17)

where \(\Lambda(x, z, t, \epsilon)\) is defined in Assumption 4.2.

We refer to the equations (4.13)–(4.16) as the reduced system. Next, we show that the reduced system (4.13)–(4.16) is a good approximation of the full system (4.7)–(4.10) as \(\epsilon\) tends to zero. In particular, we demonstrate that the error in the first and second moments between the full and the reduced system variables are \(O(\epsilon)\). To this end, we first derive some preliminary results on the moment dynamics of the full and the reduced systems.

### 4.3.2 Preliminary results

In order to quantify the error in terms of the moments, we first derive the first and second moment dynamics of the full and the reduced systems. Note that the dynamics for the stochastic fluctuations \(\psi_x\) and \(\psi_z\) in the systems (4.7) - (4.10) and (4.13)–(4.16) are linear in the variables \(\psi_x\) and \(\psi_z\). This guarantees that the resulting set moment equations for these systems will be closed.

We first present the moment dynamics of the full system (4.7) - (4.10) as follows:

**Claim 4.2.** The first and second moment dynamics for the variables \(\psi_x\) and \(\psi_z\) of the full system (4.7) - (4.10) can be expressed in the singular perturbation form

\[
\frac{d\text{E}[\psi_x]}{dt} = S_x(x, z, t)\text{E}[\psi_x] + S_z(x, z, t)\text{E}[\psi_z],
\] (4.18)

\[
\frac{d\text{E}[\psi_x\psi_x^T]}{dt} = S_x(x, z, t)\text{E}[\psi_x\psi_x^T] + S_z(x, z, t)\text{E}[\psi_z\psi_z^T] + \text{E}[\psi_x\psi_x^T]S_z(x, z, t) + (\text{E}[\psi_z\psi_x^T])^T S_z(x, z, t)^T + \sigma_z(x, z, t)\sigma_x(x, z, t)^T,
\] (4.19)
where $x$ and $z$ are the solutions of the equations (4.7) - (4.8), and the initial conditions are given by $E[\psi_x(0)] = \psi_{x0}$, $E[\psi_z(0)] = \psi_{z0}$, $E[\psi_x^T(0)] = \psi_{x0}^T$, $E[\psi_z^T(0)] = \psi_{z0}^T$.

**Proof.** See Appendix B.2. \qed

Next, we derive the moment dynamics of the reduced system (4.13)-(4.16).

**Claim 4.3.** The first and second moment dynamics for the variable $\bar{\psi}_x$ of the reduced system (4.13)-(4.14) can be written in the form

$$
\frac{dE[\bar{\psi}_x]}{dt} = S(\bar{x}, t)E[\bar{\psi}_x], \quad E[\bar{\psi}_x(0)] = \psi_{x0},
$$

and the first and second moments for the variable $\bar{\psi}_z$ of the reduced system (4.13)-(4.14) can be written in the form

$$
E[\bar{\psi}_z] = \gamma_2(\bar{x}, t)E[\bar{\psi}_x],
$$

$$
E[\bar{\psi}_z^T] = \gamma_2(\bar{x}, t)E[\bar{\psi}_x^T] + g(\bar{x}, t)g(\bar{x}, t)^T.
$$
Proof. See Appendix B.3.

Next, we derive the set of reduced-order moments obtained by setting $\epsilon = 0$ in the moment dynamics of the full system (4.18)-(4.22).

Claim 4.4. Setting $\epsilon = 0$ in the system of moment dynamics (4.18)-(4.22) and the dynamics of $x$ and $z$ given by (4.7)-(4.8), yields the moment dynamics of the reduced system (4.23)-(4.26) where the dynamics of $\bar{x}$ and $\bar{z}$ are given by (4.13) and (4.15), respectively.

Proof. See Appendix B.4.

Next, we use the results derived in this section to prove that the error in the first and second moments between the full and the reduced system variables are $O(\epsilon)$.

4.3.3 Main results

Lemma 4.1. Consider the full system in (4.7)-(4.10), the reduced system in (4.13)-(4.16), and the moment dynamics for the full and reduced systems in (4.18)-(4.22), (4.23)-(4.26) respectively. We have that, under Assumptions 4.1 - 4.3, the commutative diagram in Figure 4-1 holds.

Proof. Proof follows from Claims 4.2, 4.3 and 4.4.

Although, Lemma 4.1 shows that the setting $\epsilon = 0$ in the moment dynamics of the full system yields the moment dynamics of the reduced system, it does not guarantee that the trajectories of the moments become close to each other as $\epsilon$ decreases. Therefore, we next use the Tikhonov's theorem to prove that the moments of the reduced system are within an $O(\epsilon)$-neighborhood of the moments of the full system.

Theorem 4.1. Consider the full system (4.7)-(4.10) and the reduced system in (4.13)-(4.16). Then, under Assumptions 4.1 - 4.4, there exist $\epsilon^* > 0$ such that for $0 < \epsilon < \epsilon^*, t \in [0, t_1]$ we have

$$\|x(t) - \bar{x}(t)\| = O(\epsilon),$$

(4.27)
Figure 4-1: Setting $\epsilon = 0$ in the moment dynamics of the full system yields the moment dynamics of the reduced system.

\[
\|E[\psi_x(t)] - E[\bar{\psi}_x(t)]\| = O(\epsilon), \quad (4.28)
\]
\[
\|E[\psi_x(t)\psi_x(t)^T] - E[\bar{\psi}_x(t)\bar{\psi}_x(t)^T]\|_F = O(\epsilon), \quad (4.29)
\]

and for any $0 < t_b < t_1$ there exists $\epsilon^{**} \leq \epsilon^*$ such that for $\epsilon < \epsilon^{**}$ and $t \in [t_b, t_1]$ we have

\[
\|z(t) - \bar{z}(t)\| = O(\epsilon), \quad (4.30)
\]
\[
\|E[\psi_x(t)] - E[\bar{\psi}_x(t)]\| = O(\epsilon), \quad (4.31)
\]
\[
\|E[\psi_x(t)\psi_x(t)^T] - E[\bar{\psi}_x(t)\bar{\psi}_x(t)^T]\|_F = O(\epsilon), \quad (4.32)
\]
\[
\|E[\psi_x(t)\psi_x(t)^T] - E[\bar{\psi}_x(t)\bar{\psi}_x(t)^T]\|_F = O(\epsilon). \quad (4.33)
\]

Similar to our results in Chapter 3, the proof of this theorem (given in Appendix B.5) is based on the application of the Tikhonov's theorem to the moment dynamics of the full system which is shown to be in the singular perturbation form in Claim 4.2. Assumption 4.3 satisfies the necessary conditions on the stability of the slow
manifold of the moment equations.

Figure 4-2: Schematic diagram illustrating the model reduction approach given by Theorem 4.1. The $O(\epsilon)$-closeness of the moments of the original species concentrations and the moments obtained via the reduced system is proved in Appendix B.6.

Theorem 4.1 shows that the reduced system (4.13)- (4.16) provides a good approximation to the slow and fast variables of the full system (4.7) - (4.10) in terms of the first and second moments of the stochastic fluctuations $\psi_x$ and $\psi_z$. Thus, we have that, as $\epsilon$ tends to zero, the mean, variance and the covariance of the stochastic fluctuations $\psi_x$ and $\psi_z$ are well approximated by those of the reduced system variables.
\( \psi_x \) and \( \psi_z \). Therefore, in the case where the full system in the singular perturbation form is obtained via a coordinate change as in Claim 4.1, we have that the reduced system provides a good approximation for the variables \( v \) and \( \xi \) that consists of both slow and fast dynamics (see Appendix B.6). This is illustrated in the diagram of Figure 4-2. We also note that these results hold under the stability assumption of the Tikhonov's theorem required for the deterministic dynamics, given by Assumption 4.3, and that there are no additional stability assumptions.

From the definition of the reduced system (4.13)–(4.14), we note that the slow variable dynamics can be approximated by setting \( \epsilon = 0 \), similar to singular perturbation in deterministic systems [45]. However, the fast variable approximation \( \tilde{\psi}_z \) requires an additional term \( g(\tilde{x}, t) \). Similar to the results in Chapter 3, this additional term is required to capture the noise properties of the fast variable, which are not captured by \( \gamma_2(\tilde{x}, t) \) alone, since setting \( \epsilon = 0 \) in equation (4.10) eliminates the diffusion term \( \sigma_z(x, z, t, \epsilon) \), which contributes to the second moment of the fast variable. However, the dynamics of the slow variable \( \psi_x \) vary at a much slower rate than the dynamics of \( \psi_z \), and therefore, the noise of the fast variable given by \( \sigma_z(x, z, t, \epsilon) \) can be neglected in the slow variable approximation as it is essentially ‘filtered out’.

**Remark 1:** The stochastic fluctuations \( \psi_x \) and \( \psi_z \) are multivariate Gaussian random variables and thus their probability distributions are fully characterized by the mean and the covariance [90]. From Theorem 4.1, we have that the first and second moments of \( \psi_x(t) \) converge to the moments of the vector \( \tilde{\psi}_x(t) \), and the first and second moments of \( \psi_z(t) \) converge to the moments of \( \tilde{\psi}_z(t) \), as \( \epsilon \) tends to zero. Thus, we further have that the vectors \( \psi_x(t) \) and \( \psi_z(t) \) converge in distribution to the vectors \( \tilde{\psi}_x(t) \) and \( \tilde{\psi}_z(t) \), respectively.

Next, we investigate how the reduced system derived in this section relates to the commonly used total quasi-steady state approximation in stochastic analysis.
4.4 Validity of LNA models derived using stochastic tQSSA

The quasi-steady state approximation is widely used in the deterministic setting to obtain reduced-order models of biomolecular systems under time-scale separation conditions. Recently, the QSSA has also been extended to stochastic systems, where the fast variables are approximated by their deterministic quasi-steady state expressions in stochastic simulations. This method is termed stochastic QSSA [3, 47].

Several studies in the deterministic setting have shown that the validity of the standard QSSA is limited to certain parameter conditions. Thus, the total quasi-steady state approximation (tQSSA) has been introduced as an alternate approximation with increased accuracy [6, 12]. The tQSSA, first proposed for enzymatic reactions, considers a coordinate change to identify the slow variables in the system. This has then been adapted in the stochastic setting where stochastic simulations of the CME are performed using the deterministic quasi-steady state expressions of the fast variables obtained using the tQSSA [3, 56, 47]. There are several studies that investigate the validity of these stochastic quasi-steady state approximations. Particularly, the work by J.K Kim et al. demonstrates, via two-dimensional systems modeled by the CME and LNA, that the tQSSA provides a better approximation to the original system in comparison to the standard QSSA [47, 48]. Here, we use the results derived in Theorem 4.1 to provide a rigorous mathematical justification for the accuracy of the stochastic tQSSA for general reaction networks modeled through the LNA. The inaccuracy of the standard QSSA in the deterministic setting results from treating the species that consists of both slow and fast dynamics as purely slow variables. In contrast, the tQSSA involves defining the slow variables of the system via a coordinate change, in which the slow variable dynamics do not depend on the fast reactions. This corresponds to Claim 4.1, where a coordinate transformation is used to identify the slow and fast variables in the system. Thus, the deterministic counterpart of the reduced LNA given by equation (4.13) is equivalent to the reduced system obtained by using the tQSSA. We next derive the dynamics of the stochastic
fluctuations under the stochastic tQSSA. To this end, let \( \hat{x} \) and \( \hat{\psi}_x \) represent the variables in the LNA model obtained using the stochastic tQSSA. Then, we have that the deterministic dynamics are given by

\[
\dot{\hat{x}} = f_x(\hat{x}, \gamma_1(\hat{x}, t), t). \tag{4.34}
\]

Next, deriving the corresponding dynamics for the stochastic fluctuation using the definition of the LNA in Section 2 we have

\[
\dot{\hat{\psi}}_x = \frac{\partial f_x(\hat{x}, \gamma_1(\hat{x}, t), t)}{\partial \hat{x}} \hat{\psi}_x + \sigma_x(\hat{x}, \gamma_1(\hat{x}, t), t).
\]

Using the chain rule we can write this system as

\[
\dot{\hat{\psi}}_x = \left( \frac{\partial f_x(\hat{x}, z, t)}{\partial \hat{x}} + \frac{\partial f_x(\hat{x}, z, t)}{\partial z} \frac{\partial \gamma_1(\hat{x}, t)}{\partial \hat{x}} \right) \hat{\psi}_x \bigg|_{z=\gamma_1(\hat{x}, t)}
\]

\[+ \sigma_x(\hat{x}, \gamma_1(\hat{x}, t), t). \tag{4.35}\]

Then, by the implicit function theorem \cite{57} we have that

\[
\frac{\partial \gamma_1(\hat{x}, t)}{\partial \hat{x}} = -F_x(\hat{x}, \gamma_1(\hat{x}, t), t, 0)^{-1} \left( F_x(\hat{x}, \gamma_1(\hat{x}, t), t, 0) \right)^t,
\]

where we have from Section 4.3 that \(-F_x(\hat{x}, \gamma_1(\hat{x}, t), t, 0)^{-1} F_x(\hat{x}, \gamma_1(\hat{x}, t), t, 0) = \gamma_2(\hat{x}, t)\).

Substituting \( \partial \gamma_1(\hat{x}, t)/\partial \hat{x} = \gamma_2(\hat{x}, t) \) in (4.35), we obtain

\[
\dot{\hat{\psi}}_x = \left( S_x(\hat{x}, \gamma_1(\hat{x}, t), t) + S_x(\hat{x}, \gamma_1(\hat{x}, t), t) \gamma_2(\hat{x}, t) \right) \hat{\psi}_x
\]

\[+ \sigma_x(\hat{x}, \gamma_1(\hat{x}, t), t), \tag{4.36}\]

which is equivalent to the dynamics of \( \hat{\psi}_x \) in the reduced system (4.7) and (4.10) derived through singular perturbation. Thus, we have that the LNA model obtained under stochastic tQSSA is equivalent to the slow variable approximations...
(4.13)-(4.14) in our reduced system. This is illustrated in the schematic in Figure 4-3. Then, from Theorem 4.1 we have that the moments of the LNA model obtained using stochastic tQSSA provide a good approximation for the moments of slow variables (4.7) and (4.10) in the full system. This demonstrates that, similar to the deterministic setting, the LNA model obtained using the tQSSA provides a good approximation for the dynamics of the slow variables of the full LNA model, under all parameter conditions. Furthermore, this indicates that when the molecular counts are sufficiently large, the stochastic tQSSA yield a valid approximation for CME under time-scale separation, similar to the observations in previous studies [47, 48].
Therefore, our results provide a rigorous justification for the validity of stochastic tQSSA models, which has been absent in previous work. However, unlike the tQSSA, the reduced model derived in (4.13)-(4.16) also provide approximations for the fast variable stochastic properties. Therefore, we term our method stochastic tQSSA+. These fast variables stochastic properties can be used to analyze the reduced dynamics of the original species concentrations. Similar considerations in the literature suggests using the prefactor QSSA [4] method to study the dynamics of the fast variables in the reduced setting [47]. The prefactor QSSA method, first proposed for deterministic models of gene-regulatory networks, involves first using the coordinate transformation to identify the slow and fast variables to perform the model reduction and then transforming the reduced system variables back into the original form [4]. This is equivalent to the singular perturbation approach in deterministic setting and thus produces accurate results in deterministic models. However, J.K. Kim et al. has shown that although this provides a good approximation in deterministic models, using the reactions rates obtained through the prefactor QSSA in stochastic simulations does not produce accurate results [47, 48]. By contrast, our results provide a method to obtain accurate approximations of the original system variables.

4.5 Conclusion

In this chapter, we addressed the problem of model order reduction for biomolecular systems with time-scale separation, where the system dynamics are modeled with the LNA. We developed a reduced-order model that approximates the slow and fast dynamics of the full system when the time-scale separation is large. In particular, we showed that the error between the moment dynamics of the full system and the reduced systems is $O(\epsilon)$, where $\epsilon$ is the singular perturbation parameter that captures the time-scale separation. This error quantification only requires stability of the fast variables boundary layer dynamics, which is a simple condition to check and the same condition is required in deterministic singular perturbation. Different from existing work, we have also presented an approximation of the fast variables. Furthermore, we
show that the slow variable dynamics in our reduced-order model are equivalent to the reduced model obtained using the stochastic tQSSA approximation. Thus, our results also provide a rigorous justification for the validity of the LNA models obtained using the stochastic tQSSA. When the molecular counts are sufficiently large, this further provides a justification for the validity of the stochastic tQSSA in the CME. Since our approach, differently from the stochastic tQSSA, also provides an approximation of the fast variable stochastic properties, we have termed it the stochastic tQSSA+.
Chapter 5

Interplay between modularity and signal noise in biomolecular networks

5.1 Introduction

In this chapter we use the results obtained in Chapters 3 and 4 to analyze the stochastic effects in the modular design of biological systems. In particular, we consider several instances of network motifs that occur in gene regulatory networks and signal transduction networks in both natural and synthetic systems. We model the stochastic effects of these network motifs using reduced-order CLE or reduced-order LNA models and analyze the interplay between the conditions required for modularity and low signal noise in the networks. The results from this analysis can be used to achieve biological system designs with high accuracy and precision in signals by preserving modularity and reducing signal noise.

5.2 Modularity in gene-regulatory networks

The notion of modularity within networks has been recognized as a key feature linking biology to engineering and many studies in systems biology have focused on identifying functional modules in biological networks [1, 29, 60, 72]. Modularity also plays a fundamental role in the field of synthetic biology, where the goal is to build biological
units that can be interconnected to form networks that perform complex tasks [68].

However, it has been shown that the interconnection of modules causes a change in the pre-characterized dynamics of individual modules, making it difficult to build networks with desired characteristics. This is due to loading effects between modules that appear at interconnections, similar to loading effects in electrical circuits. In [15] Del Vecchio et. al., used singular perturbation techniques to quantify this effect and termed it 'retroactivity', which has been shown to increase with high demand from the 'load'. They also show that the effects of retroactivity could be attenuated by placing an 'insulation device' between the modules, inspired by non-inverting amplifiers in electronics.

In addition to retroactivity, another cause for loss of modularity is the competition for shared resources. As described in Chapter 2, the process of protein production requires cellular resources such as RNA polymerase and ribosomes which are shared by all transcriptional and translational processes that occur the cell. In general models of gene-regulatory networks it is typically assumed that these resources are present in abundance. However, in the case where these resources are limited, competition for them can lead to undesirable cross-talk between different proteins that do not have a direct regulatory link. This is especially important in synthetic biology where the addition of extra components can increase the demand and thus the competition for these resources.

Another source of signal perturbation in biological networks is the signal noise. One of the main methods for reducing the relative noise in a signal is to increase signal amplitude. However, it is often the case that reducing the signal noise leads to higher perturbation due to retroactivity and resource competition. Here, we analyze such trade-offs between modularity and signal noise in several network motifs that occur in both synthetic and natural networks. The network motifs we consider are interconnections of transcriptional components introduced in Chapter 2.
5.3 Trade-off between retroactivity and noise in a cascade of transcriptional components

The first network motif we consider is an interconnection of two transcriptional components shown in Figure 5-1, in which transcription factor Y activates the expression of a fluorescent protein G. Such systems are ubiquitous in synthetic genetic circuits as an indirect way of measuring the concentration of a transcription factor of interest, Y in this case. In fact, it is reasonable to think that the concentration of the fluorescent protein G should follow that of Y, possibly with some lag due to the process of gene expression encapsulated by the measuring device. Here, we study how well the concentration of G tracks that of Y in the presence of noise.

![Figure 5-1](image)

Figure 5-1: Protein X acts as an input to the upstream transcriptional component, which produces the output protein Y. The downstream transcriptional components takes protein Y as an input and produces protein G.

The chemical reactions for this system can be written as follows: $X + p_0 \xrightleftharpoons{\alpha_1}{\alpha_2} C_0, C_0 \xrightarrow{\beta_1} Y + C_0, Y \xrightarrow{\delta_1} \phi, Y + p \xrightarrow{\alpha_3}{\alpha_4} C, C \xrightarrow{\beta_2} G + C, G \xrightarrow{\delta_2} \phi [2, 14]$. Protein X binds to promoter $p_0$ and produces complex $C_0$ where $\alpha_1$ and $\alpha_2$ are the association and dissociation rate constants. $\beta_1$ is the total production rate constant of protein Y considering both transcription and translation rates. $\delta_1$ is the decay rate constant of protein Y, which includes both degradation and dilution of the protein. Similarly, $\alpha_3$ and $\alpha_4$ are the association and dissociation rate constants for protein Y and the promoter $p_0$, $\beta_2$ is the total production rate constant of protein G and $\delta_2$ is the decay rate constant of protein G. Since DNA does not dilute cell growth, the total amount of promoter in the system is conserved giving $p_{T0} = p_0 + C_0$ and $p_T = p + C [2]$. 

63
Denoting the system volume by $\Omega$, the chemical Langevin equations for the system are given by

\[
\frac{dC_0}{dt} = \alpha_1 X (pT_0 - C_0) - \alpha_2 C_0 + \sqrt{\frac{\alpha_1 X (pT_0 - C_0)}{\Omega}} \Gamma_1 - \sqrt{\frac{\alpha_2 C_0}{\Omega}} \Gamma_2,
\]

\[
\frac{dY}{dt} = \beta_1 C_0 - \delta_1 Y + \sqrt{\frac{\beta_1 C_0}{\Omega}} \Gamma_3 - \sqrt{\frac{\delta_1 Y}{\Omega}} \Gamma_4 - \alpha_3 Y (pT - C) + \alpha_4 C + \sqrt{\frac{\alpha_3 Y (pT - C)}{\Omega}} \Gamma_5 + \sqrt{\frac{\alpha_4 C}{\Omega}} \Gamma_6,
\]

\[
\frac{dC}{dt} = \alpha_3 Y (pT - C) - \alpha_4 C + \sqrt{\frac{\alpha_3 Y (pT - C)}{\Omega}} \Gamma_5 - \sqrt{\frac{\alpha_4 C}{\Omega}} \Gamma_6,
\]

\[
\frac{dG}{dt} = \beta_2 C - \delta_2 G + \sqrt{\frac{\beta_2 C}{\Omega}} \Gamma_7 - \sqrt{\frac{\delta_2 G}{\Omega}} \Gamma_8,
\]

where $\Gamma_i$ for $i = 1, \ldots, 8$ are independent Gaussian white noise processes. The binding of a transcription factor to downstream promoter sites introduces an additional rate of change in the dynamics of the transcription factor, which is represented by the boxed terms in equation (5.1) for the transcription factor $Y$. This additional rate of change, known as retroactivity, causes a change in the dynamics of the transcription factor’s concentration with respect to the isolated case, that is, when the transcription factor is not binding [15, 27]. It was also shown in the works of [15] and [27] that increasing the number of downstream binding sites $pT$ increases the effect of retroactivity on the transcription factors.

The nominal and perturbed trajectories for $Y$ and $G$ for different amounts of $pT$ can be seen in Figure 5-2. The nominal system dynamics, without perturbation due to retroactivity or noise, are obtained by simulating the ODE model obtained when $\Gamma_i = 0$ for $i = 1, \ldots, 8$ and the boxed terms are zero in the system (5.1). The perturbed trajectories are obtained using Gillespie’s direct method [21]. For lower values of $pT$ the signal $G$ closely follows the nominal signal, but the signal is highly perturbed by noise. As $pT$ increases the noise in the signal $G$ decreases, however, the signal is highly attenuated due to retroactivity. This observation is consistent with the fact that using a high gene copy number (large $pT$) is seen as a way of reducing noise in gene expression and protein production [71, 83]. However, the downside of this is that
increasing $p_T$ alters the dynamics of the input transcription factor (i.e. the protein Y), as experimentally observed in [41]. For signal Y, by contrast, both retroactivity and noise increase as $p_T$ is increased. This is consistent with prior observations in [42], where it was shown that increasing the copy number, and consequently increasing retroactivity, leads to a lower signal-to-noise ratio of transcription factors.

Figure 5-2: Nominal and perturbed signals. $G$ is obtained by simulating system (5.1) using the Gillespie algorithm [21]. The parameter values are $X = 2 + 1.5\sin(\omega t)$ nM, $\alpha_1 = 1\text{nM}^{-1}\text{s}^{-1}$, $\alpha_2 = 20\text{s}^{-1}$, $\alpha_3 = 1\text{nM}^{-1}\text{s}^{-1}$, $\alpha_4 = 100\text{s}^{-1}$, $\beta_1 = 0.01\text{s}^{-1}$, $\beta_2 = 0.1\text{s}^{-1}$, $\delta_1 = \delta_2 = 0.01\text{s}^{-1}$, $p_{T0} = 100\text{nM}$ and $\omega = 0.002\text{ rad/s}$.

Figure 5-3: The signal ‘r’ denotes the retroactivity to the upstream system. System 1 represents the nominal system in the absence of any perturbations. System 2 represents the system is perturbed only with retroactivity. System 3 represents the system perturbed with both retroactivity and intrinsic noise. $\Gamma_Y$ encapsulates the noise in the upstream component given by $\Gamma_i$ for $i = 1, \ldots, 4$ and $\Gamma_G$ encapsulates the noise in the downstream component given by $\Gamma_i$ for $i = 5, \ldots, 8$. 
In the sequel, we mathematically quantify the above trade-offs between retroactivity and noise for proteins Y and G. To this end, we formally introduce System 1 as the nominal system, System 2 as an intermediate system perturbed only with retroactivity, and System 3 as the perturbed system including both retroactivity and noise, given in Figure 5-3. Next, we derive the dynamics for each of these systems. The system (5.1) exhibits time-scale separation as the binding/unbinding reactions between transcription factors and promoter sites are much faster than protein production/decay [2]. Thus, we can represent the system dynamics in the standard singular perturbation form by defining the small parameter $\epsilon = \delta_1 / \alpha_2 \ll 1$. Representing the system variables by the non-dimensional quantities $c_0 = C_0 / p_T_0$, $y = Y / (\beta_1 p_T_0 / \delta_1)$, $c = C / p_T$, $g = G / (\beta_2 p_T / \delta_2)$, and $\bar{t} = t \delta_1$, and defining the dissociation constants $k_{d1} = \alpha_2 / \alpha_1$ and $k_{d2} = \alpha_4 / \alpha_3$ with $a = \alpha_4 / \alpha_2$, we can take the system to the standard singular perturbation form using the change of variable $v = y + p_T \delta_1 \epsilon c$, which yields

\[
\begin{align*}
\frac{dc_0}{d\bar{t}} &= \frac{X}{k_{d1}} - c_0 + \sqrt{\epsilon \frac{X}{k_{d1} p_T_0 \Omega}} \tilde{\Gamma}_1 - \sqrt{\epsilon \frac{c_0}{p_T_0 \Omega}} \tilde{\Gamma}_2, \\
\frac{dv}{d\bar{t}} &= c_0 - (v - \frac{p_T \delta_1}{\beta_1 p_T_0} c) + \sqrt{\frac{\delta_1 c_0}{\beta_1 p_T_0 \Omega}} \tilde{\Gamma}_3 - \sqrt{\frac{\delta_1 (v - \frac{p_T \delta_1}{\beta_1 p_T_0} c)}{\beta_1 p_T_0 \Omega}} \tilde{\Gamma}_4, \\
\frac{dc}{d\bar{t}} &= \frac{a \beta_1 p_T_0 (v - \frac{p_T \delta_1}{\beta_1 p_T_0} c)}{k_{d2} \delta_1} - ac + \sqrt{\frac{a \beta_1 p_T_0 (v - \frac{p_T \delta_1}{\beta_1 p_T_0} c)}{k_{d2} \delta_1 p_T \Omega}} \tilde{\Gamma}_5 - \sqrt{\frac{ac}{p_T \Omega}} \tilde{\Gamma}_6, \\
\frac{dg}{d\bar{t}} &= \frac{\delta_2}{\delta_1} c - \frac{\delta_2}{\delta_1} g + \sqrt{\frac{\delta_2^2}{\delta_1 \beta_2 p_T \Omega}} c \tilde{\Gamma}_7 - \sqrt{\frac{\delta_2^2}{\delta_1 \beta_2 p_T \Omega}} g \tilde{\Gamma}_8, 
\end{align*}
\]

where we have assumed that the binding between the proteins and promoter sites are weak, giving $C_0 \ll p_T_0$ and $C \ll p_T$, and $\tilde{\Gamma}_i$ for $i = 1, ..., 8$ represent white noise processes in the time-scale $\bar{t}$.

It follows that the system (5.2) fits the structure of the full system (4.7)-(4.8) in Chapter 3 with $v$ and $g$ as the slow variables and $c_0$ and $c$ as the fast variables. We have that the drift terms and the squared diffusion terms are linear in the state variables, satisfying Assumptions 3.1 - 3.2. The matrix $B_2$ defined in Assumption 3.2
is given by

\[
\begin{bmatrix}
-1 & 0 \\
0 & -\frac{aP}{k_d} - a
\end{bmatrix},
\]

where we have that all the parameter constants are positive. Thus, the matrix \(B_2\) is Hurwitz, satisfying Assumption 3.3. Therefore, the assumptions of Theorem 3.1 are satisfied.

We note that, due to the square-root form of the diffusion terms, the system (5.2) does not satisfy the sufficient Lipschitz continuity conditions for the existence of a unique solution for SDEs. Furthermore, the system parameters do not satisfy the conditions for the existence of a unique solution for affine SDEs in [16]. The existence of a solution for chemical Langevin equations where the arguments of the square-root diffusion terms remain positive is an ongoing research question [91, 74]. However, the validity of the chemical Langevin equation representation for chemical kinetics is based on the assumption that the molecular counts are sufficiently large [22]. In line with this, the work in [84] considers several examples of one-dimensional systems and show that the probability of molecular counts reaching zero decreases as the initial condition increases. Considering higher dimensional models, in [32], we show that the minimum time for the molecular counts to reach a lower bound starting from a given set of initial conditions increases as the initial conditions become appropriately large (as defined in [32]), thereby keeping the argument of the square-root positive for a longer time interval.

Next, setting \(\epsilon = 0\), we obtain the reduced-order system

\[
\begin{align*}
\frac{dv}{dt} &= \frac{X}{k_{d1}} - (1-R)v + \sqrt{\frac{\delta_1 X}{\beta_1 pT_0 \Omega}} \tilde{\Gamma}_3 - \sqrt{\frac{\delta_1 (1-R)v}{\beta_1 pT_0 \Omega}} \tilde{\Gamma}_4, \\
\frac{dg}{dt} &= \frac{\delta_2 \beta_1 pT_0 v}{\delta_1^2 (p_T + k_{d2})} - \frac{\delta_2 g}{\delta_1} + \sqrt{\frac{\delta_1^2 \beta_1 pT_0 v}{\delta_1^2 \beta_2 p_T (p_T + k_{d2}) \Omega}} \tilde{\Gamma}_7 - \sqrt{\frac{\delta_2^2}{\delta_1^2 \beta_2 p_T \Omega}} g \tilde{\Gamma}_8, \\
c_0 &= \frac{X}{k_{d1}} + \sqrt{\frac{X}{pT_0 k_{d1} \Omega}} N_1, \\
c &= \frac{v \beta_1 pT_0}{\delta_1 (p_T + k_{d2})} + \sqrt{\frac{v \beta_1 pT_0 k_{d2}}{\delta_1 p_T \Omega (p_T + k_{d2})^2}} N_2,
\end{align*}
\]
where \( R = \frac{p_T}{p_T + k_d} \), \( N_1 \) and \( N_2 \) are standard normal random variables. This system describes the dynamics for the perturbed system denoted by System 3 in Figure 5-3 where the dimensionless concentration for protein \( Y \) is given by \( y = v - \frac{p_T \delta_1}{\delta_1 p_T} c \). Next, the dynamics for System 2, which only includes the error due to retroactivity can be found by taking \( \Gamma_i = 0 \) for \( i = 1, \ldots, 8 \) in (5.3)–(5.6), which yields

\[
\frac{dv_R}{dt} = \frac{X}{k_{d1}} - (1 - R)v_R, \quad \frac{dg_R}{dt} = \frac{\delta_2 \beta_1 p_T v_R}{\delta_1^2 (p_T + k_{d2})} - \frac{\delta_2}{\delta_1} g_R, \quad (5.7)
\]

\[
c_{R0} = \frac{X}{k_{d1}}, \quad c_R = \frac{v_R \beta_1 p_T}{\delta_1 (p_T + k_{d2})}. \quad (5.8)
\]

Then, we can use the fast variable approximation for \( c_R \) given in (5.8) to rewrite the system dynamics in the original variable \( y_R = v_R - c_R \), to obtain

System 2: \[
y_R = (1 - R) \left( \frac{X}{k_{d1}} - y_R \right),
\]

\[
\dot{y}_R = \frac{\delta_2 \beta_1 p_T y_R}{\delta_1^2 k_{d2}} - \frac{\delta_2}{\delta_1} g_R. \quad (5.10)
\]

Similarly, the reduced-order dynamics for the nominal system (i.e without the boxed terms that represent retroactivity effects and with \( \Gamma_i = 0 \) for \( i = 1, \ldots, 8 \)) can be written as

System 1: \[
y_N = \frac{X}{k_{d1}} - y_N, \quad (5.11)
\]

\[
\dot{y}_N = \frac{\delta_2 \beta_1 p_T y_N}{\delta_1^2 k_{d2}} - \frac{\delta_2}{\delta_1} g_N. \quad (5.12)
\]

Next, using the system definitions in Figure 5-3, we define the error due to retroactivity in \( Y \) and \( G \) as \( \frac{\Delta y_N}{|y_N|} = \frac{|y_N - y_N|}{|y_N|} \) and \( \frac{\Delta g_N}{|g_N|} = \frac{|g_N - g_N|}{|g_N|} \), respectively. Similarly, the error due to noise in the signals \( Y \) and \( G \) are can be defined as \( \frac{\Delta y_R}{|y_R|} = \frac{|y_R - y_R|}{|y_R|} \) and \( \frac{\Delta g_R}{|g_R|} = \frac{|g_R - g_R|}{|g_R|} \), respectively. We consider the input \( X \) to be of the form \( X = k_1 + k_2 \sin(\omega t) \) with \( k_1 > k_2 \) to mimic a typical periodic signal from a clock [18]. As we are interested in the error in the temporal dynamics, we analyze each of the errors arising due to the time-varying component of the input \( \tilde{X} = k_2 \sin(\omega t) \).

To quantify the error due to retroactivity, we take the ratio of amplitude of the
signals $\Delta y_R$ and $\Delta g_R$ to the amplitude of the nominal signals $\Delta y_N$ and $\Delta g_N$, respectively. Therefore, the error in $y$ and $g$ due to retroactivity is given by $\frac{\Delta y_R(j\omega)}{|y_N(j\omega)|}$ and $\frac{\Delta g_R(j\omega)}{|g_N(j\omega)|}$, respectively.

To quantify the error due to noise we consider the coefficient of variation, which is a standard measure of noise, defined as the ratio of standard deviation to the mean value of a signal. Since the drift functions in the system (5.3)-(5.4) are linear, the mean signals of $y$ and $g$ are given by $y_R$ and $g_R$, respectively. Therefore, the terms $E[(\Delta y_N)^2]$ and $E[(\Delta g_N)^2]$ give the variances of signals $y$ and $g$. Then, to quantify the noise error in $Y$ we take $\sqrt{\frac{E[(\Delta y_N)^2]}{|y_R(j\omega)|^2}}$, where $|E[(\Delta y_N)^2](j\omega)|k_2$ gives the amplitude of the signal $E[(\Delta y_N)^2]$ and $|y_R(j\omega)|k_2$ gives the amplitude of the signal $y_R$ for the input $\tilde{X} = k_2 \sin(\tilde{\omega}t)$. Similarly, the noise error in $G$ can be quantified by the expression $\sqrt{\frac{E[(\Delta g_N)^2]}{|g_R(j\omega)|^2}}$.

### 5.3.1 Retroactivity error

In order to find the retroactivity error, we consider the System 1 and System 2 in Figure 5-3, for which the dynamics are given by (5.11)-(5.12) and (5.9)-(5.10). We use the linearity of the system (5.11)-(5.12) and (5.9)-(5.10) to directly evaluate the frequency response with a periodic input of the form $\tilde{X} = k_2 \sin(\tilde{\omega}t)$ and calculate the error in $Y$ and $G$ as

$$\frac{|\Delta y_R(j\omega)|}{|y_N(j\omega)|} = \frac{R\tilde{\omega}}{\sqrt{\tilde{\omega}^2 + (1 - R)^2}}, \quad (5.13)$$

$$\frac{|\Delta g_R(j\omega)|}{|g_N(j\omega)|} = \frac{R\tilde{\omega}}{\sqrt{\tilde{\omega}^2 + (1 - R)^2}}. \quad (5.14)$$

Since $R = \frac{p_T}{p_T + k_{d1}}$ monotonically increases with $p_T$, it follows that the error due to retroactivity in both $Y$ and $G$ increases as $p_T$ increases.
5.3.2 Noise error

Next, we quantify the noise error in $Y$ by considering the dynamics for System 2 and System 3 in Figure 5-3. As the drift coefficients of the system (5.3) - (5.6) are linear, we have that $E[y] = E[v] - E[c] = y_R$. Therefore, the error $E[(\Delta y_\delta)^2]$ is equivalent to the variance of $y$ given by $E[(y - E[y])^2]$. Here, we note that the dynamics of the variable $y$ consists of both slow and fast components, and therefore we require both slow and fast variable approximations to represent the dynamics of $y$ using the reduced-order model.

Thus, we use the fast variable approximation for $c$ given in (5.6) to derive the first and second moment dynamics for the variable $y$ as shown in Appendix C.1 to obtain

\[
\frac{dE[y]}{dt} = (1 - R) \left( \frac{X}{k_{d1}} - E[y] \right), \tag{5.15}
\]

\[
\frac{dE[y^2]}{dt} = (1 - R) \left[ 2 \frac{X}{k_{d1}} E[y] - 2E[y^2] + \frac{\delta_1 X}{k_{d1} p_{T0} \Omega} + \frac{\delta_1 E[y]}{\beta_1 p_{T0} \Omega} \right]. \tag{5.16}
\]

Then, using the first and second moment dynamics we find the dynamics for the variance of $y$ given by $E[(y - E[y])^2] = E[y^2] - E[y]^2$, which yields

\[
\frac{dE[(y - E[y])^2]}{dt} = (1 - R) \left[ \frac{\delta_1 X}{k_{d1} \beta_1 p_{T0} \Omega} + \frac{\delta_1 E[y]}{\beta_1 p_{T0} \Omega} - 2\delta E[(y - E[y])^2] \right], \tag{5.17}
\]

where $R = \frac{p_T}{p_T + k_{d2}}$ as defined in the derivation of the reduced system (5.3)-(5.6).

As system (5.17) is linear, we can directly evaluate its frequency response with the input $\tilde{X} = k_2 \sin(\omega t)$ which, by normalizing by the average signal $|y_R(j\omega)|k_2 = \frac{(1-R)k_2}{k_{d1}(\omega^2+(1-R)^2)}$, leads to

\[
\frac{\sqrt{E[\Delta y_\delta^2](j\omega)}}{|y_R(j\omega)|\sqrt{k_2}} = \frac{\sqrt{k_{d1} \delta_1 (\omega^2 + (1-R)^2)^{1/4}}}{\sqrt{(1-R)\beta_1 p_{T0} \Omega k_2}}
\]

Since the function $R$ monotonically increases with $p_T$, the noise error in $Y$ increases as $p_T$ increases. Therefore, decreasing the downstream copy number $p_T$ minimizes both retroactivity and noise errors in $Y$. 70
Next, we quantify the noise error in $G$ by considering the dynamics for System 2 and System 3 in Figure 5-3. Due to the linearity of the drift coefficients, the expression $\mathbb{E}[\Delta g_s^2]$, where $g_s$ was defined as $g_s = g - g_R$, gives the variance of the signal $g$. Thus, we use the dynamics of the variances of signals $v$ and $g$ to quantify the noise error in $G$. To this end, denote the variance of the signal $v$ and the covariance between $v$ and $g$ by $\mathbb{E}[(\Delta v)^2] = \mathbb{E}[(v - \mathbb{E}[v])^2]$, $\mathbb{E}[\Delta vg] = (\mathbb{E}[vg] - \mathbb{E}[g]\mathbb{E}[v])$, respectively. Then, using the moment dynamics of the system (5.3) - (5.6), the dynamics for the variances are derived as

$$
\frac{d\mathbb{E}[v]}{d\tau} = \frac{X}{k_{d1}} - (1 - R)\mathbb{E}[v],
$$

$$
\frac{d\mathbb{E}[g]}{d\tau} = \frac{\delta_2 \beta_1 P_T \mathbb{E}[v]}{\delta_1^2 (p_T + k_{d2})} - \frac{\delta_2}{\delta_1} \mathbb{E}[g],
$$

$$
\frac{d\mathbb{E}[\Delta v]^2}{d\tau} = -2(1 - R)\mathbb{E}[(\Delta v)^2] + \frac{\delta_1 X}{k_{d1} \beta_1 P_T \Omega} + \frac{\delta_1 (1 - R)\mathbb{E}[v]}{\beta_1 P_T \Omega},
$$

$$
\frac{d\mathbb{E}[\Delta vg]}{d\tau} = -((1 - R) + \frac{\delta_2}{\delta_1})\mathbb{E}[(\Delta v)g] + \frac{\delta_2 \beta_1 P_T \mathbb{E}[(\Delta v)^2]}{\delta_1^2 (p_T + k_{d2})},
$$

$$
\frac{d\mathbb{E}[\Delta g]^2}{d\tau} = 2\frac{\delta_2 \beta_1 P_T \mathbb{E}[(\Delta v)g]}{\delta_1 (p_T + k_{d2}) \mathbb{E}[(\Delta v)^2]} - 2\frac{\delta_2}{\delta_1} \mathbb{E}[\Delta g]^2 + \frac{\delta_2^2 \beta_1 P_T \mathbb{E}[v]}{\delta_1^2 \beta_2 P_T (p_T + k_{d2}) \Omega}
$$

$$
+ \frac{\delta_2}{\delta_1 \beta_2 P_T \Omega} \mathbb{E}[g],
$$

Then, evaluating the frequency response for the system (5.18)-(5.20), we can quantify the noise error in $G$ as

$$
\frac{\sqrt{\mathbb{E}[\Delta g_s^2](j\omega)}}{|g_R(j\omega)|\sqrt{k_2}} = \sqrt{\frac{\delta_2^2 \omega^2 + \delta_1^2 \omega^2}{k_2 \Omega}} \sqrt{A(p_T, \omega)}
$$

where the function $A(p_T, \omega)$ decreases with increasing $p_T$ for sufficiently small $\omega$, as shown in Appendix C.2. Therefore, as we consider an input of the form $\dot{X} = k_2 \sin(\omega t)$, where $\omega = \bar{\omega} \delta_1$, the noise error in $G$ decreases as $p_T$ increases when the input frequency $\omega$ is sufficiently smaller than the bandwidth of the nominal system given by $\delta_1$. Thus, in contrast to the noise error in $Y$, a higher value of $p_T$ should be used to decrease the noise in $G$. This is due to the fact that increasing the amount of downstream copy number $p_T$ leads to an increase in the amount of protein $G$, which
in turn reduces the amount of relative fluctuations, as observed previously [71, 83].

Furthermore, since the noise error in $Y$ increases with $p_T$ in contrast to that of $G$, and $Y$ is an input to the downstream component that produces $G$, we consider how the noise in $Y$ propagates downstream to the signal noise in $G$. To this end, we observe from Figure 5-2 that increasing $p_T$ causes an increase in the high frequency noise of signal $Y$. However, the downstream component with the output signal $G$ acts as a low-pass filter, which suggests that increasing high frequency noise content in $Y$ will have a minimal effect on the noise of $G$.

Comparing the results obtained in this section for the noise error in $G$ with the retroactivity error in $G$ given by equation (5.14) demonstrates a trade-off between stochastic and deterministic perturbation in signal $G$. Figure 5-4, illustrates this trade-off for $p_T$ in the range 1:1000 nM. Similarly, the expressions for the retroactivity error in (5.14) and the noise error in (5.21) can be used to quantify this trade-off for different parameter values and find an optimal value of $p_T$ that would minimize the combined perturbation when designing biological circuits.

![Figure 5-4: Trade-off between retroactivity and noise in signal $G$ for $p_T$ in the range 1:1000 nM (obtained using the equations (5.14) and (5.21)). The parameter values are as in Figure 4.](image)

Next, we consider another network motif that commonly occurs in gene regulatory networks due to the sharing of limited resources.
5.4 Trade-off between resource competition and noise

Figure 5-5: The proteins Y and G are expressed constitutively, in which X is a resource shared during the process of protein production.

In this section, we consider a gene regulatory network motif that occurs due to sharing of common resources. Consider the system in Figure 5-5. The proteins Y and G are produced constitutively through the promoters $p_1$ and $p_2$, respectively. We represent by X the amount of shared resources such as RNA polymerase and ribosomes used in protein production. When the shared resource X is present in a limited amount, the demand for it can introduce undesirable cross-talk between the proteins Y and G.

As seen in Section 5.3, a common method of reducing signal noise is to increase the amount of signal by increasing the promoter copy numbers $p_1$ and $p_2$. However, this would increase demand for the shared resource X and cause undesirable perturbations to signals Y and G. Here, we investigate the trade-off between decreasing the noise in protein G and increasing the perturbation on Y due to shared resource X.

The chemical reactions for the system in Figure 5-5 can be written as follows:

\[ X + p_1 \xrightarrow{k_{on1}} C_1, \quad X + p_2 \xrightarrow{k_{on2}} C_2, \]
\[ C_1 \xrightarrow{\beta_1} X + p_1 + Y, \quad C_2 \xrightarrow{\beta_2} X + p_2 + G, \]
\[ Y \xrightarrow{\delta_1} \phi, \quad G \xrightarrow{\delta_2} \phi, \]

where $C_1$ and $C_2$ represent the transcriptional components produced by the binding of resource X to $p_1$ and $p_2$, respectively. The set of reactions in (5.22) capture the
production of proteins Y and G encapsulating both transcription and translation processes, with $\beta_1$ and $\beta_2$ being the total production rate constant of protein Y and G. The decay of the proteins through degradation and dilution are shown in (5.24), in which $\delta_1$ and $\delta_2$ are decay rates. The total amount of resource X and the promoters $p_1$ and $p_2$ are conserved. Therefore, we also have the conservation laws $X_t = X + C_1 + C_2$, $p_{T1} = p_1 + C_2$, $p_{T2} = p_2 + C_2$.

Using these reactions we can write the following chemical Langevin equations for this system.

\[
\begin{align*}
\dot{Y} &= \beta_1 C_1 - \delta_1 Y + \sqrt{\beta_1 C_1 \Gamma_1} - \sqrt{\delta_1 Y \Gamma_2}, \\
\dot{G} &= \beta_2 C_2 - \delta_2 G + \sqrt{\beta_2 C_2 \Gamma_3} - \sqrt{\delta_2 G \Gamma_4}, \\
\dot{C}_1 &= k_{\text{on}1}(p_{T1} - C_1)(X_t - C_1 - C_2) - k_{\text{off}1} C_1 - \beta_1 C_1 \quad + \sqrt{k_{\text{on}1}(p_{T1} - C_1)(X_t - C_1 - C_2) \Gamma_5} - \sqrt{k_{\text{off}1} C_1 \Gamma_6} - \sqrt{\beta_1 C_1 \Gamma_7}, \\
\dot{C}_2 &= k_{\text{on}2}(p_{T2} - C_2)(X_t - C_1 - C_2) - k_{\text{off}2} C_2 - \beta_2 C_2 \quad + \sqrt{k_{\text{on}2}(p_{T2} - C_2)(X_t - C_1 - C_2) \Gamma_8} - \sqrt{k_{\text{off}2} C_2 \Gamma_9} - \sqrt{\beta_2 C_2 \Gamma_{10}},
\end{align*}
\]

in which $\Gamma_i$ for $i = 1, \ldots, 10$ are white noise processes and we have assumed that the system volume $\Omega = 1 \text{ nM}^{-1}$. Figure 5-6, shows a simulation of the proteins Y and G for different values of $p_{T2}$, performed using the Gillespie algorithm. It can be seen that the increase of promoter concentrations $p_{T2}$ cause a drop in the steady state concentration of Y, and this perturbation increases with increasing $p_{T2}$. This perturbation is due to the competition for shared resource X as also experimentally observed in several studies [28, 69]. However, in order to decrease the amount of relative noise on the signal G it is desirable to have a high amount of promoter concentrations $p_{T2}$. In the sequel, we mathematically quantify the error on G due to noise and the error on Y due to resource competition.

The binding/unbinding reactions are much faster than protein production and decay, and thus we can define a small constant $\epsilon = \delta_1/k_{\text{off}1} \ll 1$. Then denote by $k_{d1} = k_{\text{off}1}/k_{\text{on}1}$ and $k_{d2} = k_{\text{off}2}/k_{\text{on}2}$ the dissociation constants and let $a = k_{\text{off}1}/k_{\text{off}2}$. 

74
Figure 5-6: Concentrations of Y and G obtained by simulating system (5.25) using the Gillespie algorithm [21]. The parameter values are $X = 10\text{nM}$, $\beta_1 = \beta_2 = 0.1\text{s}^{-1}$, $\delta_1 = \delta_2 = 0.01\text{s}^{-1}$, $p_{T1} = 200\text{nM}$, $k_{d1} = k_{d2} = 100\text{nM}$.

Using these, we can then write the system in the singular perturbation form:

$$
\dot{Y} = \beta_1 C_1 - \delta_1 Y + \sqrt{\beta_1} C_1 \Gamma_1 - \sqrt{\delta_1} Y \Gamma_2
$$

$$
\dot{G} = \beta_2 C_2 - \delta_2 G + \sqrt{\beta_2} C_2 \Gamma_3 - \sqrt{\delta_2} G \Gamma_4
$$

$$
\epsilon \dot{C}_1 = \frac{\delta_1}{k_{d1}} p_{T1}(X_t - C_1 - C_2) - \delta_1 C_1 - \epsilon \beta_1 C_1 + \sqrt{\frac{\epsilon \delta_1}{k_{d1}}} p_{T1}(X_t - C_1 - C_2) \Gamma_5 - \sqrt{\epsilon \delta_1} C_1 \Gamma_6 - \epsilon \sqrt{\beta_1} C_1 \Gamma_7,
$$

$$
\epsilon \dot{C}_2 = \frac{a \delta_1}{k_{d2}} p_{T2}(X_t - C_1 - C_2) - a \delta_1 C_2 - \epsilon \beta_2 C_2 + \sqrt{\frac{\epsilon a \delta_1}{k_{d2}}} p_{T2}(X_t - C_1 - C_2) \Gamma_8 - \sqrt{\epsilon a \delta_1} C_2 \Gamma_9 - \epsilon \sqrt{\beta_2} C_2 \Gamma_{10},
$$

where we have assumed that the binding between the promoter sites and the resource X is weak, resource X is weak, giving rise to $p_{T1} \ll C_1$ and $p_{T2} \ll C_2$. We have that the proteins Y and G are the slow variables in the system and the complexes $C_1$ and $C_2$ are the fast variables. This system fits the structure of the full systems for CLEs in singular perturbation form with Y and G being the slow variables and $C_1$ and $C_2$ being the fast variables. Assumptions 3.1 and 3.2 are satisfied since drift and diffusion terms are linear in the state variables. For Assumption 3.3, we compute the
matrix $B_2$ which yields

$$B_2 = \begin{bmatrix} -\delta_1 p_{T1}/k_{d1} - \delta_1 & -\delta_1 p_{T1}/k_{d1} \\ -a\delta_1 p_{T2}/k_{d2} & -a\delta_1 p_{T2}/k_{d2} - a\delta_1 \end{bmatrix}.$$  

The eigenvalues of matrix $B_2$ are given by $-\delta_1$ and $-\frac{\delta_1 (k_{d2} p_{T1} + k_{d1} (k_{d2} + p_{T2}))}{k_{d1} k_{d2}}$, in which the parameters $\delta_1, k_{d1}, p_{T2}, k_{d2}, p_{T2}$ are positive. Thus, the matrix $B_2$ is Hurwitz, which satisfies Assumption 3.3. Then, setting $\epsilon = 0$, we obtain the reduced system

$$\begin{align*}
\dot{Y} &= \frac{\beta_1 X_t p_{T1}/k_{d1}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1} - \delta_1 Y + \sqrt{\frac{\beta_1 X_t p_{T1}/k_{d1}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1}} \Gamma_1 - \sqrt{\delta_1 Y} \Gamma_2 \quad (5.26) \\
\dot{G} &= \frac{\beta_2 p_{T2} X_t/k_{d2}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1} - \delta_2 G + \sqrt{\frac{\beta_2 p_{T2} X_t/k_{d2}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1}} \Gamma_3 - \sqrt{\delta_2 G \Gamma_4}. \quad (5.27)
\end{align*}$$

for the slow variables $Y$ and $G$.

In order to quantify the error in protein $Y$ due to resource competition, we introduce a nominal system for protein $Y$ in the absence of competition for resources by protein $G$. This is obtained by setting $p_{T2} = 0$ in (5.26) - (5.27).

$$\frac{dY_N}{dt} = \frac{\beta_1 X_t p_{T1}/k_{d1}}{p_{T1}/k_{d1} + 1} - \delta_1 Y_N + \sqrt{\frac{\beta_1 X_t p_{T1}/k_{d1}}{p_{T1}/k_{d1} + 1}} \Gamma_1 - \sqrt{\delta_1 Y_N} \Gamma_2. \quad (5.28)$$

We then proceed to compute the steady state errors in $Y$ and $G$ due to resource competition and noise. We use the moment dynamics of the perturbed and nominal signals to compute the errors in signals $Y$ and $G$. To quantify the error on $Y$ due to resource competition, we use the relative error between the signals $Y$ and $Y_N$. To quantify the error in $G$ due to noise we use the coefficient of variation, which is defined as the ratio of standard deviation to the mean.
5.4.1 Error in Y due to resource competition

To quantify the error due to resource competition on Y, we derive the moment dynamics of the perturbed and the nominal signals Y and \( Y_N \), respectively.

\[
\begin{align*}
\frac{d\mathbb{E}[Y]}{dt} &= \frac{\beta_1 X_t p_{T1}/k_{d1}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1} - \delta_1 \mathbb{E}[Y], \\
\frac{d\mathbb{E}[Y_N]}{dt} &= \frac{\beta_1 X_t p_{T1}/k_{d1}}{p_{T1}/k_{d1} + 1} - \delta_1 \mathbb{E}[Y_N].
\end{align*}
\]

Solving for the steady states for the perturbed and nominal systems we obtain

\[
\begin{align*}
\mathbb{E}[Y] &= \frac{\beta_1 X_t p_{T1}/k_{d1}}{\delta_1 (p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1)}, \\
\mathbb{E}[Y_N] &= \frac{\beta_1 X_t p_{T1}/k_{d1}}{\delta_1 (p_{T1}/k_{d1} + 1)}.
\end{align*}
\]

Then, we can find the relative error in Y due to resource competition as

\[
\frac{\mathbb{E}[Y_N] - \mathbb{E}[Y]}{\mathbb{E}[Y_N]} = \frac{p_{T2}/k_{d2}}{(p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1)}
\]

(5.29)

It can be seen that the error in Y due resource competition increases as \( p_{T2} \), the promoter corresponding to protein G, increases. Thus, increasing the amount of the protein G increases the error in protein Y.

5.4.2 Noise error in G

To quantify the noise error in G, we first compute the first and second moment dynamics for G given by

\[
\begin{align*}
\frac{d\mathbb{E}[G]}{dt} &= \frac{\beta_2 p_{T2} X_t/k_{d2}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1} - \delta_2 \mathbb{E}[G], \\
\frac{d\mathbb{E}[G^2]}{dt} &= 2\frac{\beta_2 p_{T2} X_t/k_{d2}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1} \mathbb{E}[G] - 2\delta_2 \mathbb{E}[G^2] + \frac{\beta_2 p_{T2} X_t/k_{d2}}{p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1} + \delta_2 \mathbb{E}[G].
\end{align*}
\]

Then computing the steady state of these dynamics we can find the steady state...
coefficient of variation of $G$ to be

$$\frac{\sqrt{\mathbb{E}[G^2] - \mathbb{E}[G]^2}}{\mathbb{E}[G]} = \sqrt{\frac{\delta_2(p_{T2}/k_{d2} + p_{T1}/k_{d1} + 1)}{\beta_2 p_{T2} X_{t}/k_{d2}}}.$$  \tag{5.30}

It can be seen that the coefficient of variation of $G$ decreases as $p_{T2}$ increases, in contrast to the error in $Y$ due to resource competition. Therefore, the expressions (5.29) and (5.30) quantify the trade-off between decreasing noise in protein $G$ and increasing the error in $Y$ due to resource competition. This trade-off is illustrated in Figure 5-7, for a set of parameter values. These expressions can be used to identify optimal promoter concentrations that causes the minimum perturbations in designing biological systems.

![Figure 5-7: Trade-off between reducing noise in $G$ and increasing the error due to resource competition in $Y$ for $p_T$ in the range 50:1000 nM (obtained using the equations (5.29) and (5.30)). The parameter values are $X = 10$ nM, $k_{d1} = 100$ nM, $\beta_2 = 0.01$ s$^{-1}$, $\delta_2 = 0.01$ s$^{-1}$, $p_{T1} = 100$ nM.](image)

5.5 Noise buffering through decoy sites

In this section, we consider another network motif that is typically seen in gene-regulatory networks. Consider the system given in Figure 5-8, where the transcription
factor X binds to the promoter $p_2$ and regulates the production of protein G, while also binding to a non-regulatory binding site $p_1$. This type of non-regulatory binding sites - referred to as decoy sites - occur widely in gene-regulatory network [55]. As we have discussed in Section 5.2, downstream binding sites cause a perturbation to the upstream signal due to retroactivity, causing a lag in the response of signal X. The experimental observations of these effects can be found in [41, 61].

However, it has been suggested that decoy sites, such as the promoter sites $p_1$ in Figure 5-8 can attenuate the steady state noise of the target protein G [8]. Similar stochastic effects have also been studied in [79], using the chemical Master equation. Here, we use the chemical Langevin equation to analyze the effects of decoy sites $p_1$ on the steady state noise of the signals X and G.

![Diagram of transcription factor X regulating protein G](image)

Figure 5-8: Transcription factor X regulates the production of protein G, while also binding to non-regulatory binding site $p_1$.

The chemical reactions for the system can be written as follows: $\phi \xrightarrow{k_{\delta_1}} X$, $X + p_1 \xrightarrow{k_{on1}} C_1$, $X + p_2 \xrightarrow{k_{on2}} C_2$, $C_2 \xrightarrow{\beta} C_2 + G$, $G \xrightarrow{\delta} \phi$, where $k$ is the production rate of X, $k_{on1}, k_{off1}$ and $k_{on2}, k_{off2}$ are the binding/unbinding rate constants between the transcription factor X and the promoters $p_1$ and $p_2$, $\beta$ is the production rate of the protein G, and $\delta_1, \delta_2$ are the decay rate constants of X and G respectively, which includes both degradation and dilution. The total amount of each promoter is conserved, and hence we can write $p_{T1} = p_1 + C_1$ and $p_{T2} = p_2 + C_2$. Denote by $\Omega$ the cell volume, and let $\Omega = 1 \text{ nM}^{-1}$ for simplicity. Then, the chemical Langevin
equations for the system can be written as

\[
\frac{dX}{dt} = k - \delta_1 X - k_{on1}X(p_{T1} - C_1) + k_{off1}C_1 - k_{on2}X(p_{T2} - C_2) + k_{off2}C_2
\]
\[
+ \sqrt{k\Gamma_1} - \sqrt{\delta_1 X \Gamma_2} - \sqrt{k_{on1}X(p_{T1} - C_1)\Gamma_3} - \sqrt{k_{off1}C_1\Gamma_4}
\]
\[
- \sqrt{k_{on2}X(p_{T2} - C_2)\Gamma_5} - \sqrt{k_{off2}C_2\Gamma_6}.
\]

\[
\frac{dC_1}{dt} = k_{on1}X(p_{T1} - C_1) - k_{off1}C_1 + \sqrt{k_{on1}X(p_{T1} - C_1)\Gamma_3} - \sqrt{k_{off1}C_1\Gamma_4}, \quad (5.31)
\]
\[
\frac{dC_2}{dt} = k_{on2}X(p_{T2} - C_2) - k_{off2}C_2 + \sqrt{k_{on2}X(p_{T2} - C_2)\Gamma_5} - \sqrt{k_{off2}C_2\Gamma_6},
\]
\[
\frac{dG}{dt} = \beta C_2 - \delta_2 G + \sqrt{\beta C_2\Gamma_7} - \sqrt{\delta_2 G\Gamma_8},
\]

where \( \Gamma_i \) are independent white noise processes. We assume that the binding between the transcription factor \( X \) and the promoters are weak, giving \( p_{T1} \gg C_1 \) and \( p_{T2} \gg C_2 \). Therefore, we can write the system (5.31) in the form

\[
\frac{dX}{dt} = k - \delta_1 X - k_{on1}Xp_{T1} + k_{off1}C_1 - k_{on2}Xp_{T2} + k_{off2}C_2
\]
\[
+ \sqrt{k\Gamma_1} - \sqrt{\delta_1 X \Gamma_2} - \sqrt{k_{on1}Xp_{T1}\Gamma_3} - \sqrt{k_{off1}C_1\Gamma_4}
\]
\[
- \sqrt{k_{on2}Xp_{T2}\Gamma_5} - \sqrt{k_{off2}C_2\Gamma_6},
\]

\[
\frac{dC_1}{dt} = k_{on1}Xp_{T1} - k_{off1}C_1 + \sqrt{k_{on1}Xp_{T1}\Gamma_3} - \sqrt{k_{off1}C_1\Gamma_4},
\]
\[
\frac{dC_2}{dt} = k_{on2}Xp_{T2} - k_{off2}C_2 + \sqrt{k_{on2}Xp_{T2}\Gamma_5} - \sqrt{k_{off2}C_2\Gamma_6},
\]
\[
\frac{dG}{dt} = \beta C_2 - \delta_2 G + \sqrt{\beta C_2\Gamma_7} - \sqrt{\delta_2 G\Gamma_8}.
\]

We have that the binding/unbinding reactions are much faster than protein production and decay [2], and thus we can write \( \epsilon = \delta_1/k_{off1} \), where \( \epsilon \ll 1 \). Letting \( k_{d1} = k_{off1}/k_{on1}, k_{d2} = k_{off2}/k_{on2} \), and \( a = k_{off2}/k_{off1} \) we have \( k_{on1} = \delta_1/(\epsilon k_{d1}), k_{on2} = a\delta_1/(\epsilon k_{d2}), k_{off1} = \delta_1/\epsilon, \) and \( k_{off2} = a\delta_1/\epsilon \). Then, with the change of variable \( y = X + C_1 + C_2 \), we can take the system into standard singular perturbation form

\[
\frac{dy}{dt} = k - \delta_1(y - C_1 - C_2) + \sqrt{k\Gamma_1} - \sqrt{\delta_1(y - C_1 - C_2)\Gamma_2}, \quad (5.32)
\]
\[
\frac{dC_1}{dt} = \frac{\delta_1}{k_{d1}}(y - C_1 - C_2)p_{T1} - \delta_1C_1 + \sqrt{\epsilon} \frac{\delta_1}{k_{d1}}(y - C_1 - C_2)p_{T1}\Gamma_3 - \sqrt{\epsilon}\frac{\delta_1C_1\Gamma_4}{k_{d1}},
\] (5.33)

\[
\frac{dC_2}{dt} = \frac{a\delta_1}{k_{d2}}(y - C_1 - C_2)p_{T2} - a\delta_1C_2 + \sqrt{\epsilon} \frac{a\delta_1}{k_{d2}}(y - C_1 - C_2)p_{T2}\Gamma_5 - \sqrt{\epsilon}a\delta_1C_2\Gamma_6,
\] (5.34)

\[
\frac{dG}{dt} = \beta C_2 - \delta_2G + \sqrt{\beta C_2\Gamma_7} - \sqrt{\delta_2\Gamma_8}. 
\] (5.35)

Similar to the previous systems in Sections 5.3 - 5.4, we have that this system fits the structure of the full systems for CLEs in singular perturbation form with \(y\) and \(G\) being the slow variables and \(C_1\) and \(C_2\) being the fast variables. We have that the Assumptions 3.1 and 3.2 are satisfied since drift and diffusion terms are linear in the state variables. Considering Assumption 3.3, we compute the matrix \(B_2\) which yields

\[
B_2 = \begin{bmatrix}
-\delta_1p_{T1}/k_{d1} - \delta_1 & -\delta_1p_{T1}/k_{d1} \\
-a\delta_1p_{T2}/k_{d2} & -a\delta_1p_{T2}/k_{d2} - a\delta_1
\end{bmatrix}.
\]

The eigenvalues of \(B_2\) are given by \(-\delta_1\) and \(-\delta_1\frac{k_{d2}p_{T1} + k_{d1}(k_{d2} + p_{T2})}{k_{d1}k_{d2}}\) where the parameters \(\delta_1, k_{d1}, p_{T2}, k_{d2}, p_{T2}\) are positive. Therefore, we have that the matrix \(B_2\) is Hurwitz, satisfying Assumption 3.3.

Then, setting \(\epsilon = 0\), we obtain the function

\[
C_1 = \frac{p_{T1}/k_{d1}y}{p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1},
\]

\[
C_2 = \frac{p_{T2}/k_{d2}y}{p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1}.
\]

Then, to obtain the function \(g(y, t)\) we consider the equation

\[
g(y, t)g(y, t)^T B_2^T + B_2g(y, t)g(y, t)^T = -\Lambda(y, \gamma_1(y, t), t, 0),
\] (5.36)
where $\Lambda$ is given by

$$
\Lambda = \begin{bmatrix}
\frac{2\delta_1 p T_1/k_{d1} y}{p T_1/k_{d1} + p T_2/k_{d2} + 1} & 0 \\
0 & \frac{2\delta_1 p T_2/k_{d2} y}{p T_1/k_{d1} + p T_2/k_{d2} + 1}
\end{bmatrix}.
$$

Then, solving the set of linear equations in (5.36), we find that the matrix $g(y, t)g(y, t)^T$ is given by

$$
\frac{1}{(1 + p T_1/k_{d1} + p T_2/k_{d2})^2} \begin{bmatrix}
p T_1/k_{d1}(1 + p T_2/k_{d2}) y & -p T_1/k_{d1} p T_2/k_{d2} y \\
-p T_1/k_{d1} p T_2/k_{d2} y & p T_2/k_{d2}(1 + p T_1/k_{d1}) y
\end{bmatrix}
$$

and therefore we have

$$
g(y, t) = \frac{1}{\sqrt{S + \Delta}} \begin{bmatrix}
p T_1/k_{d1}(1 + p T_2/k_{d2}) y + \sqrt{\Delta} & -p T_1/k_{d1} p T_2/k_{d2} y \\
-p T_1/k_{d1} p T_2/k_{d2} y & p T_2/k_{d2}(1 + p T_1/k_{d1}) y + \sqrt{\Delta}
\end{bmatrix},
$$

where $S = (p T_1/k_{d1} + p T_2/k_{d2} + 2 p T_1/k_{d1} p T_2/k_{d2}) y$ and $\Delta = p T_1/k_{d1} p T_2/k_{d2}(1 + p T_1/k_{d1} + p T_2/k_{d2}) y^2$.

Then, the reduced system is given by

$$
\frac{d y}{d t} = k - \delta_1 y \left(\frac{1}{p T_1/k_{d1} + p T_2/k_{d2} + 1} + \sqrt{k} \Gamma_1 - \sqrt{\delta_1 y \left(\frac{1}{p T_1/k_{d1} + p T_2/k_{d2} + 1}\right)} \Gamma_2\right),
$$

$$
\frac{dG}{dt} = \beta p T_2/k_{d2} y - \delta_2 G + \sqrt{\frac{\beta p T_2/k_{d2} y}{p T_1/k_{d1} + p T_2/k_{d2} + 1}} \Gamma_7 - \sqrt{\delta_2 G \Gamma_8},
$$

$$
C_1 = \frac{p T_1/k_{d1} y}{p T_1/k_{d1} + p T_2/k_{d2} + 1}
+ \frac{1}{\sqrt{S + \Delta}} \left(\frac{(p T_1/k_{d1}(1 + p T_2/k_{d2}) y + \sqrt{\Delta})}{(1 + p T_1/k_{d1} + p T_2/k_{d2})} N_1 - \frac{p T_1/k_{d1} p T_2/k_{d2} y}{(1 + p T_1/k_{d1} + p T_2/k_{d2})} N_2\right),
$$

(5.37) (5.38) (5.39)
$C_2 = \frac{pT_2/k_d y}{pT_1/k_d1 + pT_2/k_d2 + 1}$

$$+ \frac{1}{\sqrt{S + \sqrt{\Delta}}} \left( -\frac{pT_1/k_d pT_2/k_d y}{(1 + pT_1/k_d1 + pT_2/k_d2)} N_1 + \frac{(pT_2/k_d(1 + pT_2/k_d2)y + \sqrt{\Delta}) N_2}{(1 + pT_1/k_d1 + pT_2/k_d2)} \right),$$

(5.40)

where $N_1$ and $N_2$ are standard normal random variables.\(^\text{1}\)

Next, we analyze how the steady state noise of the proteins X and G vary with increasing amount of decoy sites $pT_1$. As a measure of noise we use the coefficient of variation, which is defined as the ratio of standard deviation to the mean, which can be computed using the moment dynamics of the reduced system.

### 5.5.1 Noise analysis

We use the moment dynamics of the system (5.37) - (5.40) to analyze the effect increasing $pT_1$ on noise of the signals X and G. Towards this end, we first compute the first and second order moment dynamics of the slow variables $y$ and $G$ as

$$\frac{dE[y]}{dt} = k - \frac{\delta_1}{pT_1/k_d1 + pT_2/k_d2 + 1} E[y],$$

$$\frac{dE[G]}{dt} = \frac{\beta pT_2/k_d}{pT_1/k_d1 + pT_2/k_d2 + 1} E[y] - \delta_2 E[G],$$

$$\frac{dE[y^2]}{dt} = 2kE[y] - \frac{2\delta_1}{pT_1/k_d1 + pT_2/k_d2 + 1} E[y^2] + k + \frac{\delta_1}{pT_1/k_d1 + pT_2/k_d2 + 1} E[y],$$

$$\frac{dE[G^2]}{dt} = \frac{2\beta pT_2/k_d}{pT_1/k_d1 + pT_2/k_d2 + 1} E[yG] - 2\delta_2 E[G^2] + \frac{\beta pT_2/k_d}{pT_1/k_d1 + pT_2/k_d2 + 1} E[y] + \delta_2 E[G],$$

$$\frac{dE[yG]}{dt} = kE[G] - \frac{\delta_1}{pT_1/k_d1 + pT_2/k_d2 + 1} E[yG] + \frac{\beta pT_2/k_d}{pT_1/k_d1 + pT_2/k_d2 + 1} E[y^2] - \delta_2 E[yG].$$

(5.41)

Computing the steady state of this system, we can derive an expression for the steady state coefficient of variation of signal $G$ as

$$\frac{\sqrt{E[G^2] - E[G]^2}}{E[G]} = \sqrt{\frac{\delta_1 \delta_2 (\delta_1 + \beta pT_2/k_d2 + \delta_2 (1 + pT_1/k_d1 + pT_2/k_d2))}{\beta k pT_2/k_d2 (\delta_1 + \delta_2 (1 + pT_1/k_d1 + pT_2/k_d2))}}$$

\(^1\text{We further illustrate the convergence of the moments of the reduced system to those of the full system via simulations as shown in Appendix C.4.1.}\)
From this expression it can be seen that increasing the number of decoy sites \( p_{T1} \) leads to a decrease in the noise of the protein \( G \). Thus, as observed in previous studies [79], it follows that higher number of decoy sites can be used to attenuate the steady state signal noise in the downstream target protein \( G \).

Next, we analyze the noise on protein \( X \). As we obtained the reduced-order model by using a coordinate transformation \( y = X + C_1 + C_2 \), we now require the fast variable approximations to derive an expression for the noise on protein \( X \). Thus, using the approximations (5.39)-(5.40), we can derive the moment dynamics for \( C_1 \) and \( C_2 \) as

\[
\begin{align*}
E[C_1^2] &= \frac{p_{T1}/k_{d1}y^2}{p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1}, \\
E[C_2^2] &= \frac{p_{T2}/k_{d2}y^2}{p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1}, \\
E[C_1]\ &= \frac{(p_{T1}/k_{d1}y)^2}{(p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1)^2} + \frac{p_{T1}/k_{d1}(1 + p_{T2}/k_{d2})y}{(1 + p_{T1}/k_{d1} + p_{T2}/k_{d2})^2}, \\
E[C_2]\ &= \frac{(p_{T2}/k_{d2}y)^2}{(p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1)^2} + \frac{p_{T2}/k_{d2}(1 + p_{T1}/k_{d1})y}{(1 + p_{T1}/k_{d1} + p_{T2}/k_{d2})^2}, \\
E[C_1C_2]\ &= \frac{p_{T1}/k_{d1}p_{T2}/k_{d2}y^2}{(p_{T1}/k_{d1} + p_{T2}/k_{d2} + 1)^2} - \frac{p_{T1}/k_{d1}p_{T2}/k_{d2}y}{(1 + p_{T1}/k_{d1} + p_{T2}/k_{d2})^2}.
\end{align*}
\]

Using these moments and the moments of \( y \) from the system (5.41), we find that

\[
\sqrt{\frac{E[X^2] - E[X]^2}{E[X]}} = \sqrt{\frac{k}{\delta_1}}
\]

Interestingly, the coefficient of variation of protein \( X \) at steady state does not depend on the decoy sites \( p_{T1} \). Similar results have been observed previously in a cascase of transcriptional components studied by Ghaemi and Del Vecchio [20] where it is shown that the upstream and the downstream system are statistically independent at steady state, and also in the work by Soltani et al. in [79]. This then leads to the question of how the noise in \( G \) decreases due to decoy sites while the coefficient of variance of \( X \) remains constant. An explanation to this can be found by analyzing the power
spectrum of the signal $X$. Through simulations, we observe that increasing the decoy sites shift signal content of $X$ to high frequency (Figure 5-9). The transcriptional component producing $G$ acts as a low pass filter selecting only the low frequency components from $X$. Thus, the noise in $G$ decreases while there is no change of the variance of signal $X$.

![Figure 5-9: Increasing decoy sites $p_1$ shifts the noise content of signal $X$ to high frequency. Simulation performed with Gillespie algorithm. The parameters used are $k = 1\text{nM}\text{s}^{-1}$, $\delta_1 = 0.1\text{s}^{-1}$, $\delta_1 = 0.1\text{s}^{-1}$, $k_{on1} = 1\text{s}^{-1}$, $k_{off1} = 100\text{nM}\text{s}^{-1}$, $k_{on2} = 1\text{s}^{-1}$, $k_{off2} = 100\text{nM}\text{s}^{-1}$, $p_T = 1\text{nM}$, $\beta = 10\text{s}^{-1}$](image)

5.6 Insulation through phosphorylation cycles

In [15, 40], Del Vecchio et al. show that placing an ‘insulation device’ between two components is a way of attenuating retroactivity between the components. The insulation device proposed in their work uses high-gain negative feedback, and can be realized in biological systems by using phosphorylation cycle with high substrate and phosphatase concentrations. Experimental realizations of such devices are seen in [63, 61]. In this section, we consider the effects of attenuating retroactivity through insulation on the noise of the signals in the system.

Consider the network motif shown in Figure 5-10. Protein X is phosphorylated by kinase Z and produces protein $X^*$, which is dephosphorylated by phosphatase Y. The phosphorylated protein $X^*$ binds to the promoter $p$ and produces protein $G$. 

85
Having high substrate and phosphatase concentrations X and Y will attenuate the retroactivity to the input Z. Here, we study how the noise of the proteins X* and G change with increasing substrate and phosphatase concentrations.

We model the system dynamics using the Linear Noise Approximation. Towards this end we first consider the chemical reactions for the system in Figure 5-10. The phosphorylation and dephosphorylation processes can be modeled as one-step reactions [14], which yields

\[ X + Z \xrightarrow{k_1} X^* + Z, \quad X^* + Y \xrightarrow{k_2} X + Y, \]

where \( k_1 \) and \( k_2 \) are the phosphorylation and dephosphorylation reaction rate constants. The binding between the protein X* and promoter p produces complex a complex C, giving

\[ X^* + p \xrightarrow{k_{on}} \frac{k_{on}}{k_{off}} C, \]

where \( k_{on} \) and \( k_{off} \) are the binding and unbinding rate constants. Then, the production and decay of protein G is given by

\[ C \xrightarrow{\beta} C + G, \quad G \xrightarrow{\delta} \phi, \]

where \( \beta \) is the production rate constant that includes both transcription and translation and \( \delta \) is the decay rate constant, which includes both dilution and degradation of the protein. The total concentration of protein X and promoter p are conserved, giving \( X_T = x + x^* + x \) and \( p_T = p + c \), where we use the lower-case letters to denote
the macroscopic concentrations. Then, the macroscopic reaction rate equations for this system can be written as

\[
\frac{dx^*}{dt} = k_1 Z(t)(X_T - x^* - c) - k_2 Y x^* - k_{on} x^*(P_T - c) + k_{off} c,
\]

\[
\frac{dc}{dt} = k_{on} x^*(P_T - c) - k_{off} c,
\]

\[
\frac{dg}{dt} = \beta c - \delta g.
\]

In this analysis, we assume that the total concentration of protein X is much larger than the total concentration of promoter p, giving \(X_T \gg P_T\), which yields

\[
\frac{dx^*}{dt} = k_1 Z(t)(X_T - x^*) - k_2 Y x^* - k_{on} x^*(P_T - c) + k_{off} c, \tag{5.42}
\]

\[
\frac{dc}{dt} = k_{on} x^*(P_T - c) - k_{off} c, \tag{5.43}
\]

As binding and unbinding reactions are much faster than phosphorylation and de-phosphorylation reactions, we have that \(k_{off} \gg k_2 Y\). Thus, we can define the small parameter \(\epsilon = k_2 Y / k_{off}\). Then, letting \(k_d = k_{off} / k_{on}\) be the dissociation constant for the binding reaction between \(X^*\) and \(p\), the system (5.42) - (5.43), can be written as

\[
\frac{dx^*}{dt} = k_1 Z(t)(X_{tot} - x^*) - k_2 Y x^* - \frac{k_2 Y}{\epsilon k_d} x^*(P_{tot} - c) + \frac{k_2 Y}{\epsilon} c, \tag{5.44}
\]

\[
\frac{dc}{dt} = \frac{k_2 Y}{\epsilon k_d} x^*(P_{tot} - c) - \frac{k_2 Y}{\epsilon} c, \tag{5.45}
\]

\[
\frac{dg}{dt} = \beta c - \delta g. \tag{5.46}
\]

Although, the singular perturbation parameter appears in the system of equations (5.44) - (5.46), we note that the slow and fast dynamics are not well separated and
the system not in the standard singular perturbation form given in (4.7) - (4.8). Thus, we consider the change of coordinates \( v = x^* + c \), which yields

\[
\begin{align*}
\frac{dv}{dt} &= k_1 Z(t)(X_T - v + c) - k_2 Y(v - c), \\
\frac{dg}{dt} &= \beta c - \delta g, \\
\frac{dc}{dt} &= k_2 Y(v - c)(p_T - c) - k_2 Yc.
\end{align*}
\]

(5.47) (5.48) (5.49)

where we have that the slow variables are \( v \) and \( g \), and the fast variable is given by \( c \). Considering Claim 4.1, we have that the coordinate change \( v = x^* + c \) corresponds to \( A_x = [1 \ 1 \ 0; 0 \ 0 \ 1]^T \), \( A_z = [0 \ 1 \ 0] \), and \( y = [x^*, \ c, \ g]^T \), \( x = [v, g]^T \) and \( z = c \). Therefore, we can write the following equations for the dynamics of the stochastic fluctuations.

\[
\begin{align*}
\frac{d\psi_v}{dt} &= (-k_1 Z(t) - k_2 Y)\psi_v + (k_1 Z(t) + k_2 Y)\psi_c \\
&\quad + \sqrt{k_1 Z(t)(X_T - v + c)\Gamma_1 - k_2 Y(v - c)\Gamma_2}, \\
\frac{d\psi_g}{dt} &= \beta \psi_v - \delta \psi_g + \sqrt{\beta c \Gamma_3 - \delta g \Gamma_4}, \\
\frac{d\psi_c}{dt} &= \frac{k_2 Y(p_T - c)}{k_d} \psi_v \\
&\quad + \left(-\frac{k_2 Y p_T}{k_d} - \frac{k_2 Y v}{k_d} + \frac{k_2 Y}{k_d} 2c - k_2 Y\right) \psi_c \\
&\quad + \sqrt{\epsilon \frac{k_2 Y}{k_d} (v - c)(p_T - c)\Gamma_5 - \sqrt{\epsilon k_2 Y c \Gamma_6}.
\end{align*}
\]

(5.50) (5.51) (5.52)

Next, we derive the reduced-order dynamics for the system (5.47) - (5.49) using Theorem 4.1. From the equations (5.47) - (5.49), it follows that the system functions are polynomials of the state variables. Therefore, the Assumptions 4.1 and 4.2 are satisfied. Evaluating \( f_z = \frac{k_2 Y}{k_d} (v - z)(p_T - z) - k_2 Yz = 0 \), yields the unique solution 

\[
\gamma_1(v) = \frac{1}{2}(v + p_T + k_d) - \frac{1}{2}\sqrt{(v + p_T + k_d)^2 - 4vp_T},
\]

feasible under the physical constraints \( 0 \leq c \leq p_T \). We note that \( \frac{df_z}{dz} \) is negative for all parameter values and thus Assumption 4.3 is satisfied. Next, to determine the fast variable approximation for
\( \psi_c \) which is in the form of equation (4.16), we write

\[
\left( -\frac{k_2 Y p_T}{k_d} - \frac{k_2 Y v + k_2 Y}{k_d} - 2c - k_2 Y \right) g(y)g(y)^T + g(y)g(y)^T \left( -\frac{k_2 Y p_T}{k_d} - \frac{k_2 Y v}{k_d} + \frac{k_2 Y}{k_d} 2c - k_2 Y \right) = \frac{k_2 Y}{k_d} (v - c)(p_T - c) - k_2 Y c
\]

which yields

\[
g(y) = \sqrt{\frac{k_2 Y (v - c)(p_{tot} - c) + k_2 Y c}{2k_2 Y k_d (v + p_{tot} + k_d)^2 - 4vp_{tot}}}.
\]

Then, the reduced system is given by

\[
\frac{dv}{dt} = k_1 Z(t)(X_T - v + c) - k_2 Y (v - c), \quad (5.53)
\]

\[
\frac{dg}{dt} = \beta c - \delta g, \quad (5.54)
\]

\[
\frac{d\psi_v}{dt} = (-k_1 Z(t) - k_2 Y)\psi_v + \frac{(k_1 Z(t) + k_2 Y)(p_T - c)}{(p_T + v - 2c + k_d)} \psi_v
\]

\[+ \sqrt{k_1 Z(t)(X_T - v + c)}\Gamma_1 - \sqrt{k_2 Y (v - c)}\Gamma_2, \quad (5.55)
\]

\[
\frac{d\psi_g}{dt} = \beta \frac{(p_T - c)}{(p_T + v - 2c + k_d)} \psi_v - \delta \psi_g + \sqrt{\beta c \Gamma_3 - \sqrt{\delta \Gamma_4}}, \quad (5.56)
\]

\[
c = \frac{1}{2} (v + p_T + k_d) - \frac{1}{2} \sqrt{(v + p_T + k_d)^2 - 4p_T v} \quad (5.57)
\]

\[
\psi_c = \frac{(p_T - c)}{(p_T + v - 2c + k_d)} \psi_v
\]

\[
+ \sqrt{\frac{k_2 Y (v - c)(p_{tot} - c) + k_2 Y c}{2k_2 Y k_d (v + p_{tot} + k_d)^2 - 4vp_{tot}}} N(0, 1), \quad (5.58)
\]

where \( N(0, 1) \) is a standard normal random variable. From Theorem 4.1, we have

that the error between the moments of the reduced and the full system are of \( O(\epsilon) \).

In the sequel, we analyze how the noise properties of the proteins \( X^* \) and \( G \) vary as the total substrate concentration \( X_T \) and phosphatase concentration \( Y \) increases. As

\[\text{We further illustrate the convergence of the moments of the reduced system to those of the full system through simulations as shown in Appendix C.4.2.}\]
a measure of noise intensity we use the coefficient of variation, which is defined as the ratio of standard deviation to the mean. To compute the mean and the variance of the signals we use the moment dynamics of the system (5.53) - (5.57). We first consider a constant of the form \( Z(t) = k \) and analytically compute the steady state coefficient of variations of the proteins \( X^* \) and \( G \). We then consider a time-varying input of the form \( Z(t) = k + m \sin(\omega t) \) and numerically analyze the change in coefficient of variation of proteins \( X^* \) and \( G \) for varying \( X_T \) and \( Y \).

### 5.6.1 Noise analysis

We first derive the first and second moment dynamics of the system (5.53) - (5.58), using that \( \Gamma_1 - \Gamma_4 \) are white noise processes and \( N(0,1) \) is a normal random variable with zero mean and variance 1. This yields

\[
\frac{dv}{dt} = k_1 Z(t)(X_T - v + c) - k_2 Y(v - c), \tag{5.59}
\]

\[
\frac{dg}{dt} = \beta c - \delta g, \tag{5.60}
\]

\[
\frac{d\text{E}[\psi_v]}{dt} = (-k_1 Z(t) - k_2 Y)\text{E}[\psi_v] + \frac{(k_1 Z(t) + k_2 Y)(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v], \tag{5.61}
\]

\[
\frac{d\text{E}[\psi_g]}{dt} = \frac{\beta(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v] - 2\delta \text{E}[\psi_g], \tag{5.62}
\]

\[
\frac{d\text{E}[\psi_v^2]}{dt} = 2(-k_1 Z(t) - k_2 Y)\text{E}[\psi_v^2] + \frac{2(k_1 Z(t) + k_2 Y)(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v^2]
\]

\[
+ k_1 Z(t)(X_T - v + c) + k_2 Y(v - c), \tag{5.63}
\]

\[
\frac{d\text{E}[\psi_v \psi_g]}{dt} = (-k_1 Z(t) - k_2 Y)\text{E}[\psi_v \psi_g] + \frac{(k_1 Z(t) + k_2 Y)(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v \psi_g]
\]

\[
+ \frac{\beta(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v^2] - \delta \text{E}[\psi_v \psi_g], \tag{5.64}
\]

\[
\frac{d\text{E}[\psi_g^2]}{dt} = \frac{2\beta(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v \psi_g] - 2\delta \text{E}[\psi_g^2] + \beta c + \delta g, \tag{5.65}
\]

\[
\text{E}[\psi_c] = \frac{(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v], \tag{5.66}
\]

\[
\text{E}[\psi_v \psi_v] = \frac{(p_T - c)}{(p_T + v - 2c + k_d)} \text{E}[\psi_v^2]. \tag{5.67}
\]
\[ E[\psi_c^2] = \left( \frac{p_T - c}{p_T + v - 2c + k_d} \right)^2 E[\psi_v^2] + \frac{k_2 Y (v - c)(p_{tot} - c) + k_2 Y c}{2k_2 Y \sqrt{(v + p_{tot} + k_d)^2 - 4vp_{tot}}} \]  \hfill (5.68)

Using zero initial conditions for the stochastic fluctuations, it follows from the equations (5.61) - (5.62) that \( E[\psi_v] = 0 \) and \( E[\psi_g] = 0 \) for all time. Thus, considering the definition of the Linear Noise Approximation in (4.1) - (4.2), we have that the mean and the variance of concentration of protein G is given by \( g \) and \( E[\psi_g^2] \). Therefore, the coefficient of variation for protein G can be calculated as \( \sqrt{\frac{E[\psi_g^2]}{g}} \). Similarly, for protein \( X^* \), we have that the mean value of the concentration is given by \( x^* \) where \( x^* = v - c \), and the variance is given by \( E[\psi_x^2] \) where \( \psi_x = \psi_v - \psi_c \). Thus, we can use the moments of fast variable approximations given in (5.66) - (5.68), to calculate the coefficient of variation of \( X^* \) in the form \( \sqrt{\frac{E[\psi_x^2]}{x^*}} \).

**Steady state analysis**

In this section, we consider a constant input of the form \( Z(t) = k \) and analytically compute the coefficient of variation for proteins G and \( X^* \). For ease of analysis, we represent the deterministic dynamics in terms of the original variables \( x^* \) and \( g \) using the fast variable approximation \( c \). We have that \( x^* = v - c \), which yields the dynamics

\[ \frac{dx^*}{dt} = \frac{dv}{dt} - \frac{dc}{dt}. \]

Then, by using the chain rule we obtain

\[ \frac{dx^*}{dt} = \frac{1}{1 + \frac{dc}{dx^*}} \frac{dv}{dt}. \]  \hfill (5.69)

Next, by using \( v = x^* + c \) in (5.57) we can rewrite the fast variable approximation for \( c \) in terms of \( x^* \) as

\[ c = \frac{x^* p_T}{(x^* + k_d)}. \]  \hfill (5.70)
Then, using the equation (5.69), (5.70) and \( v = x^* + c \) in the set of moment equations (5.59) - (5.68), we have

\[
\frac{dx^*}{dt} = \frac{1}{(x^* + k_d)^2 + prk_d} (k_1Z(t)(X_T - x^*) - k_2Yx^*),
\]

(5.71)

\[
\frac{dg}{dt} = \beta \frac{x^*p_T}{(x^* + k_d)} - \delta g,
\]

(5.72)

\[
\frac{dE[\psi_v^2]}{dt} = 2(-k_1Z(t) - k_2Y)E[\psi_v^2] + \frac{2(k_1Z(t) + k_2Y)k_dpr}{(prk_d + (x^* + k_d)^2)}E[\psi_v^2]
\]

+ \( k_1Z(t)(X_T - x^*) + k_2Y(v - \frac{x^*p_T}{(x^* + k_d)}) \),

(5.73)

\[
\frac{dE[\psi_v\psi_g]}{dt} = (-k_1Z(t) - k_2Y)E[\psi_v\psi_g] + (k_1Z(t) + k_2Y)k_dprE[\psi_v\psi_g]
\]

+ \( \frac{\beta k_dpr}{(prk_d + (x^* + k_d)^2)}E[\psi_v^2] - \delta E[\psi_v\psi_g] \),

(5.74)

\[
\frac{dE[\psi_g^2]}{dt} = \frac{2\beta k_dpr}{(prk_d + (x^* + k_d)^2)}E[\psi_v\psi_g] - 2\delta E[\psi_g^2] + \beta c + \delta g,
\]

(5.75)

\[
E[\psi_v] = \frac{k_dpr}{(prk_d + (x^* + k_d)^2)}E[\psi_v],
\]

(5.76)

\[
E[\psi_c] = \left( \frac{k_dpr}{(prk_d + (x^* + k_d)^2)} \right)^2 E[\psi_v^2] + \frac{x^*p_Tk_d}{kdpr + (x^* + k_d)^2}.
\]

(5.77)

Then, by taking \( X_T = aY \) and computing the steady state value of the above set of equations (5.71) - (5.77), we have that the coefficient of variation of the protein G is given by

\[
\frac{\sqrt{E[\psi_g^2]}}{g} = \sqrt{A(Y)},
\]

where the function \( A(Y) \) has a negative derivative with respect to \( Y \) for all parameter conditions, as shown in Appendix C.3. Thus, the steady state noise in protein G decreases with increasing concentrations of \( Y \) and \( X_T \).

Next, we use the moments of the fast variable approximation \( \psi_c \) to compute the noise in protein X. We have that \( \psi_x = \psi_v - \psi_c \), which yields \( E[\psi_x^2] = E[\psi_v^2] - 2E[\psi_v\psi_c] + \)
\( \mathbb{E}[\psi_2^2] \). Therefore, using the expressions (5.77) and (5.57), we find that

\[
\mathbb{E}[\psi_2^2] = \left( \frac{(x^* + k_d)^2}{(pT_k_d + (x^* + k_d)^2)} \right)^2 \mathbb{E}[\psi_2^2] + \frac{x^* pTk_d}{k_d pT + (x^* + k_d)^2}.
\] (5.78)

Then, using the steady state value of \( \mathbb{E}[\psi_2^2] \) from (5.71), we have that the coefficient of variation of \( X^* \) is given by

\[
\frac{\sqrt{\mathbb{E}[\psi_2^2]}}{x} = \sqrt{B(Y)}
\]

where the function \( B(Y) \) has a negative derivative with respect to \( Y \) for all parameter conditions, as shown in Appendix C.3. Thus, similar to protein \( G \) the steady state noise in protein \( X^* \) decreases as \( Y \) and \( X_T \) increases.

Therefore, it can be seen that increasing the substrate and phosphatase concentrations also decrease the steady state noise in the proteins \( X^* \) and \( G \), in addition to attenuating retroactivity.

**Time-varying analysis.**

Next, we investigate the noise properties of the proteins \( G \) and \( X^* \) for a time-varying input \( Z(t) = 5 + 3.75(\omega t) \) nM where \( \omega = 0.02 \) rad/min. By simulating the moment equations (5.59) - (5.60) and (5.63) - (5.68), we numerically compute the time-varying coefficient of variation for proteins \( X^* \) and \( G \) for varying values of \( X_T \) and \( Y \) in the range 50nM to 1000nM. We then compare the maximum coefficient of variation of the permanent response (i.e. after a transient of 700 minutes has elapsed) for different values of \( X_T \) and \( Y \). The results are shown in Figure 5-11.

It can be seen from Figure 5-11 that for this choice of parameters, the maximum CV of protein \( X^* \) increases as the concentrations of \( X_T \) and \( Y \) increase, in contrast to the previous section. However, for protein \( G \) the maximum CV decreases with increasing \( X_T \) and \( Y \), and therefore, the increasing CV of \( X^* \) does not have an effect on the noise of protein \( G \). A reason for this can be that the high-gain feedback shifts the noise of \( X^* \) to high frequency, and as the downstream component acts as
Figure 5-11: Maximum coefficient of variation for proteins $G$ and $X^*$. The phosphorylation and dephosphorylation rate constants were approximated from Chen et. al [10] as $k_1 = 0.0005 \text{ nMmin}^{-1}$ and $k_2 = 0.0005 \text{ (nMmin)^{-1}}$. The remaining parameters were chosen as $\beta = 0.5 \text{ min}^{-1}$, $\delta = 0.01 \text{ min}^{-1}$, $p_T = 5 \text{ nM}, k_d = 10\text{nM}$ to be in the nominal range for bacterial cells [27].

...a low-pass filter it only selects the low frequency components of $X^*$, thus reducing the noise of the signal $G$. Further analysis of this system could be used to point towards optimal concentrations of substrate and phosphatase concentrations that give maximum accuracy in signal transmission.

Our analysis in this section shows that, in contract to the previous network motifs that demonstrated trade-offs between modularity and signal noise, insulation mechanism based on high-gain feedback could be used to build networks that attenuate both retroactivity and noise leading to accurate signal transmission in networks.

5.7 Conclusion

In this chapter, we analyzed the interplay between modularity and signal noise in several network motifs that occur in synthetic and systems biology. It was seen that often there are trade-offs between ensuring modularity of the network and reducing the noise in the signals. However, we also find that insulation devices designed to attenuate retroactivity through high-gain negative feedback can also attenuate the sig-
nal noise at steady state. These findings indicate that phosphorylation cycles, which make up signaling networks in natural systems, may also play a role in attenuating noise and ensuring accurate signal transmission in nature.

The analysis of the stochastic properties in this study were facilitated by using the model order reduction techniques developed in Chapters 3 and 4 to obtain simplified models. Often, the species of interest included mixed variables, where both slow and fast variable approximations were used to compute the noise properties.
Chapter 6

Reduced-order LNA models for gene-regulatory networks

6.1 Introduction

In this chapter, we utilize the reduced-order model for Linear Noise Approximation developed in Chapter 4 to obtain reduced-order models for stochastic dynamics in gene-regulatory networks. The stochastic QSSA, the process of using the deterministic QSSA to reduce the deterministic system is a convenient way to obtain a reduced LNA similar to the QSSA approximation in deterministic systems. However, the validity of this reduction remains elusive. As we previously discussed in Section 4.4, several researchers have investigated the validity of this approximation in the context of the CME and LNA. However, their analyses are limited simulation studies and two-dimensional systems.

Here, we use the model reduction approach developed in Chapter 4, the stochastic tQSSA+, to obtain a set of reduced moment dynamics for gene-regulatory networks. These moments dynamics accurately capture the moments of the original species dynamics, and therefore, can be used to investigate the validity of stochastic QSSA in LNA models of gene-regulatory networks. Furthermore, the reduced moment dynamics resulting from the stochastic tQSSA+ can be directly used to analyze the statistical properties of the system under all parameter conditions. In the sequel, we
begin by introducing the structure of the gene-regulatory network considered. We then present the LNA model using the stochastic QSSA approximation and compute its moments. Next, we present the moment dynamics computed using the stochastic tQSSA+ approach. Comparing these two sets of moment dynamics we identify conditions under which the LNA model based on the stochastic QSSA provides a good approximation.

6.2 Network structure and problem description

We consider a gene regulatory network with \( N \) nodes, where each node is a transcriptional component as shown in Figure 6-1a. Let \( Y_j \) denote the output protein species produced by each node \( j \). We denote by \( \eta_j \) the number of input TFs that node \( j \) can take, and we assume that the binding between input TFs and promoter sites are independent. i.e. each node \( j \) has \( \eta_j \) promoter binding sites to which TFs can bind independently. The binding between an input TF and promoter produces a promoter-TF complex, which in turn leads to protein production. In the case of independent binding as considered in this work, the binding between input TFs and the promoter binding sites produces \( 2^{\eta_j} \) distinct promoter-TF complexes for each node \( j \). For example, a node with two inputs will have 4 distinct promoter-TF complexes as shown in Figure 6-1b. Denote by \( D_{jk} \) be the \( k^{th} \) promoter-TF complex species in node \( j \), where \( k = 1, \ldots, 2^{\eta_j} \).

We next define the following sets: Let \( \mathcal{I}_j \) be the set of input TFs to node \( j \), and \( \mathcal{D}_j \) be the set of complexes in node \( j \). Let \( \Theta_{jk} \) be the set of input TFs that are bound to \( D_{jk} \). Then, we have the following reactions for node \( j \):

For \( Y_i \in \mathcal{I}_j \) and \( D_{jk} \in \mathcal{D}_j|Y_i \notin \Theta_{jk} \), we have the binding reaction

\[
nY_i + D_{jk} \xrightarrow{\alpha_{ji}} D_{jk},
\]

where \( \Theta_{ji} = D_{jk} \cup Y_j \). The parameters \( \alpha_{ji} \) and \( \beta_{ji} \) denote the binding and unbinding rate constants between the TF \( Y_i \) and the corresponding promoter site of node \( j \).
(a) Gene-regulatory network where each node is a transcriptional component. The promoter region in the transcriptional component contains binding sites for each input TF for the transcriptional component.

(b) The possible promoter-TF complexes due to independent binding in a transcriptional component with two inputs. i.e. \( \eta_j = 2 \). The promoter region consists of two binding sites corresponding to each input. The TFs can bind to their respective binding site independent of whether the other binding site is occupied. (Other binding types such as cooperative and competitive binding are not considered in this work.)

Figure 6-1: (a) Gene-regulatory network. (b) Promoter-TF complexes for a node with two inputs, where each input binds independently as considered in this work.

Next, we consider the production of protein through the promoter-TF complexes. The process of protein production occurs in two steps: transcription, where the DNA sequence is transcribed into mRNA by RNA polymerase and translation, where the mRNA is used to produce the proteins. Here, we lump these processes together and denote the protein production by one reaction:

\[
D_{jk} \xrightarrow{\pi_{jk}} D_{jk} + Y_j,
\]

where the parameter \( \pi_{jk} \) is the production rate that corresponds to the complex \( D_{jk} \). Thus, for transcriptionally inactive complexes we have \( \pi_{jk} = 0 \). We then consider the
decay of $Y_j$ through due to degradation and dilution:

$$Y_j \xrightarrow{\delta_j} \phi$$

Furthermore, we have that the total amount of promoter in a node is conserved. Therefore, we can write the following conservation law: $\sum q_j D_{jk} = P_{Tj}$, where $P_{Tj}$ denotes the total amount of promoter.

### 6.2.1 Problem description

In order to model the stochastic dynamics of this network, we can use the LNA to provide a description for the concentrations of proteins $Y_j$ and complexes $D_{ik}$ in the form

$$Y_j = y_j + \frac{\xi_j}{\sqrt{\Omega}},$$

and

$$D_j = d_{ik} + \frac{\Upsilon_{ik}}{\sqrt{\Omega}},$$

where $y_j$ and $d_{ik}$ are the deterministic trajectories, and $\chi_j$ and $\Upsilon_{ik}$ accounts for the fluctuations around the deterministic trajectory, and $\Omega$ is the system volume.

The binding and unbinding reactions occur at a much faster time-scale than the protein production and decay reactions, and therefore, this property often used in deterministic system models to provide a reduced-order approximation for the dynamics of $y_j$ by taking the complexes $d_{ik}$ to be at their quasi-steady state. The stochastic QSSA approach extends this to stochastic models, whereby the stochastic fluctuations $\xi_j$ are directly derived using the reduced-order approximation for the deterministic dynamics of $y_j$. This provides an LNA model in terms of the Hill functions that are commonly used to model transcriptional regulation in gene networks. However, the validity of this approximation is still under investigation by different researchers.

On the other hand, in Chapter 4, we introduced a model reduction approach, termed stochastic tQSSA$^+$, which provides a reduced-order LNA that approximates the system dynamics under time-scale separation conditions. Therefore, here we use
the reduced-order LNA model developed in Chapter 4 to derive a set of reduced-order representation for the moment dynamics of $y_j$ and $\xi_j$, which provide a good approximation for the moments of original variables. Thus, these moment dynamics can be use to investigate the validity of the LNA model obtained through stochastic QSSA, by comparison of the corresponding moment dynamics of the stochastic QSSA. In the following, we present the LNA model obtained using the stochastic QSSA approximation and its corresponding moments dynamics. Then, we present the moments dynamics of the LNA using stochastic $tQSSA^+$ approximation. We use these moment dynamics to investigate the conditions under which the stochastic QSSA provides a good approximation for gene-regulatory network models.

6.3 Results

In this section, we first present the LNA model based on the stochastic QSSA, which we refer to as the Hill function model, and derive the corresponding moment dynamics. We then present the reduced-order moment dynamics based on the stochastic $tQSSA^+$. Next, we use these moment dynamics to identify conditions under which the Hill function models become a good approximation of the original system dynamics. The detailed derivations of the models are presented in the Section 6.5.

6.3.1 Hill function model

The LNA model based on the stochastic QSSA is derived by approximating the concentrations of the promoter-TF complexes with their quasi-steady state expressions. With this approximation, we have the Hill function model

$$
\dot{y}_j = H_j(\tilde{u}_j) - \delta_j \tilde{y}_j, \\
\dot{\xi}_j = \sum_{q \in S_j} \frac{\partial H(\tilde{u}_j)}{\partial \tilde{y}_q} \tilde{\xi}_q - \delta_j \xi_j + \sqrt{H(\tilde{u}_j)} \Gamma_{j1} - \sqrt{\delta_j \tilde{y}_j} \Gamma_{j2}.
$$

(6.1)  
(6.2)
in which $\mathbf{u}_j = [u_{j_1}, \ldots, u_{j_n}]$ is the vector of inputs to node $j$, and $H_j(\mathbf{u}_j)$ denotes the commonly Hill function that describes the rate of production of protein $Y_j$ [2, 14]. The detailed derivation of this model and the definition of the Hill function are given in Section 6.5.

Then, using these dynamics we can derive the variance and the covariance of the fluctuations (6.2) as follows:

$$\frac{d\mathbb{E}[\xi_j^2]}{dt} = 2 \sum_{q \in \mathcal{J}_j} \frac{\partial H(\mathbf{u}_j)}{\partial \tilde{Y}_q} \mathbb{E}[\xi_q \tilde{\xi}_j] - 2\delta_j \mathbb{E}[\xi_j^2] + H_j(\mathbf{u}_j) + \delta_j \tilde{Y}_j, \quad (6.3)$$

$$\frac{d\mathbb{E}[\xi_j \xi_r]}{dt} = (-\delta_j - \delta_r) \mathbb{E}[\tilde{\xi}_j \tilde{\xi}_r] + \sum_{q \in \mathcal{J}_j} \frac{\partial H(\mathbf{u}_j)}{\partial \tilde{Y}_q} \mathbb{E}[\tilde{\xi}_q \tilde{\xi}_j] + \sum_{q \in \mathcal{J}_r} \frac{\partial H(\mathbf{u}_j)}{\partial \tilde{Y}_q} \mathbb{E}[\tilde{\xi}_q \tilde{\xi}_r], \text{ for } j \neq r. \quad (6.4)$$

Next, we present the moment dynamics of the system obtained using the stochastic tQSSA$^+$ approach.

### 6.3.2 Moments of the stochastic tQSSA$^+$ model

The derivation of the Hill function model is based on approximating the promoter-TF complexes at the quasi-steady state, and directly substituting these expressions in the dynamics of $y_j$. However, in deriving an approximation using the stochastic tQSSA$^+$, based on the singular perturbation approach, we note that $y_i$ participates in both slow and fast reactions, and thus are not the slow variables in the system. Therefore, as detailed in Section 6.5, we first use a coordinate change to identify the slow and fast variables of the system and obtain a reduced-order approximation for these using the stochastic tQSSA$^+$. We then use these approximations to derive the moments of the 'mixed' variables $y_j$ as follows:

$$\dot{y}_j = \frac{1}{1 + \mathcal{R}_j(y_j)} (H(\mathbf{u}_j) - \delta_j y_j),$$

$$\frac{d\mathbb{E}[\xi_j^2]}{dt} = \frac{1}{1 + \mathcal{R}_j(y_j)} \left[ -2\delta_j (1 + [F_j(y_j)]) \mathbb{E}[\xi_j^2] + 2 \sum_{q \in \mathcal{J}_j} \frac{\partial H_j(\mathbf{u}_j)}{\partial \tilde{Y}_q} \mathbb{E}[\xi_q \xi_j] \right]$$
\[ + H_j(u_j) \left( 1 + \left[ G_j(y_j) \right] J_j(y_j) \right) \delta_j \] + \left( 1 - \left[ G_j(y_j) \right] \right) J_j(y_j), \tag{6.5} \]

\[
\frac{dE[\xi_j \xi_r]}{dt} = \frac{1}{1 + \left[ R_j(y_j) \right]} \sum_{q \in s_j} \frac{\partial H_j(u_j)}{\partial y_q} E[\xi_q \xi_r] \sum_{q \in s_r} \frac{\partial H_r(u_r)}{\partial y_q} E[\xi_q \xi_r] - \left( 1 \right) \frac{1}{1 + \left[ R_j(y_j) \right]} (1 - \left[ F_j(y_j) \right]) E[\xi_j \xi_r] - I_{jr}(y_j, y_r),
\]

where we use \( \mathcal{E}_j \) to denote the set of nodes that are regulated by \( y_j \), the term

\[ R_{jl}(y_j) = \frac{n}{\left( pr_j/k_{dji} \right) y_j^{n-1}}, \]

quantifies the retroactivity to node \( j \) from from node \( l \) (further definition in Section 6.5), and

\[
R_j(y_j) = \sum_{l \in \mathcal{E}_j} R_{jl}(y_j), \quad F_j(y_j) = \frac{1}{\delta_j} \sum_{l \in \mathcal{E}_j} \frac{dR_{jl}(y_j)}{dy_j} \frac{dy_j}{dt},
\]

\[
G_j(y_j) = \frac{2 \sum_{l \in \mathcal{E}_j} R_{jl}(y_j) + 1}{\left( 1 + \sum_{l \in \mathcal{E}_j} R_{jl}(y_j) \right)^2} y_j,
\]

\[
J_j(y_j) = 2 \frac{\partial H_j(u_j)}{\partial y_j} \sum_{l \in \mathcal{E}_j} R_{jl}(y_l) y_j,
\]

\[
I_{jr}(y_j, y_r) = \frac{1}{1 + \sum_{l \in \mathcal{E}_j} R_{jl}(y_j)} \frac{\partial H_j(u_j)}{\partial y_r} \sum_{l \in \mathcal{E}_j} R_{rl}(y_l) y_r
\]

The boxed terms in (6.32) highlight the extra terms compared to the moment dynamics of the Hill function model (6.4). From this, it follows that the Hill function model accurately approximates the original system dynamics only when the boxed terms are sufficiently small.

Next, we investigate parameter conditions that allows the boxed terms to be small.
Conditions for validity of the Hill function model\(^1\)

From the dynamics of \(y_j\) we have that

\[
y_j(t) \leq \max\{y_j(0), H(u_j)/\delta_j\}
\]

The hill function typically takes the form

\[
H_j(u_j) = p_{Tj} \frac{\pi_{j1} + \pi_{j2} \frac{y_j}{k_{dji}} + \pi_{j3} \frac{yk}{k_{djk}} + \pi_{j4} \frac{yk}{k_{dji} k_{djk}}}{1 + \frac{y_j}{k_{dji}} + \frac{y_k}{k_{djk}} + \frac{yk}{k_{dji} k_{djk}}},
\]

\[
\leq p_{Tj} (\pi_{j1} + \pi_{j2} + \pi_{j3} + \pi_{j4}),
\]

for two inputs \(y_k, y_i\) and can be generalized to higher dimensions [27, 2].

Therefore, considering the terms

\[
R_{jl}(y_j) = n \frac{(PTl/k_{dlj})y_j^{n-1}}{(y_j/k_{dlj} + 1)^2},
\]

\[
\frac{dR_{jl}(y_j)}{dy_j} = n(p_{Tl}/k_{dlj})y_j^{n-1} \frac{-(n + 1)y_j^{n-1}/k_{dlj} + (n - 1)y_j^{-1}}{(y_j/k_{dlj} + 1)^3},
\]

we have that for given values of \(p_{Tj}, \delta_j\) and \(\pi_{jk}\) the terms \(R_{jl}(y_j)\) and \(\frac{dR_{jl}(y_j)}{dy_j}\) tend to zero as \(p_{Tl}/k_{dlj}\) tends to zero. Similarly, it can be seen that the functions \(F_j\), \(G_j\), \(J_j\) and \(I_{jr}\) are functions of \(y_j, y_r\) and the corresponding Hill functions and retroactivity terms. Therefore, we have that, for given values of \(p_{Tj}, p_{Tr}, \delta_j, \delta_r\) and \(\pi_{jk}, \pi_{rk}\), the boxed terms \(R_j, F_j, G_j, J_j\) and \(I_{jr}\) become close to zero as the ratios \(p_{Tl}/k_{dl}, p_{Tr}/k_{dr}\) tend to zero. This can be achieved by having very large dissociation constants \(k_{dlj}, k_{dlr}\). Thus, having large dissociation constants, i.e., \(p_{Tl}/k_{dlj} \ll 0\), provide a sufficient condition for the validity of the Hill function models in the case of bounded protein concentrations \(y_j\) and Hill functions \(H_j(u_j)\).

We illustrate this via simulation with an example of an incoherent feedforward network motif [2, 14] shown in Figure 6-2. We derive the LNA models for this system using the stochastic QSSA and stochastic tQSSA\(^+\) (shown in Appendix D.3) and

\(^1\)This analysis was performed jointly with Simone Bruno.
compute the sample moments by simulating these LNA models. The sample moments for this system are shown in Figures 6-3 and 6-4\(^2\). It can be seen that for large dissociation constants the moments of the Hill function model becomes close to the moments of the stochastic tQSSA\(^+\) model.

![Diagram](attachment:image.png)

Figure 6-2: Protein Y\(_1\) activates the production of protein Y\(_2\) and Y\(_3\), while Y\(_2\) represses the production of protein Y\(_3\).

### 6.4 Discussion and future work

The terms \(R_{ji}\) in the dynamics (6.32) denote the retroactivity terms that appear in the previous studies on modularity in gene-regulatory networks. In particular, as also discussed in Section 5, Del Vecchio et al. [15, 27, 14] show that the terms of the form \(R_{ji}\) physically arise due to the 'loading' that promoter binding sites apply to their transcription factor regulators. These effects, cause a change in the dynamics of transcription factors upon binding to the DNA promoter sites and are not captured by Hill function based models alone. Thus, the results obtained in this chapter indicate that, similar to deterministic systems, retroactivity effects also impact the stochastic dynamics of the system.

As previously described in Section 4.4, the use of the deterministic QSSA in stochastic models has been previously studied by several researchers in the context of CME and LNA. In particular, the study by J.K. Kim et al. considers two dimensional

\(^2\)The simulations in these figures were performed by Simone Bruno.
Figure 6-3: Sample variances computed using the Hill function based LNA model and the stochastic tQSSA+ based LNA model. Sample variances are computed using 150,000 simulation runs. The parameter values are $\pi_{10} = 0.01 \text{ min}^{-1}$, $\pi_{21} = 1 \text{ min}^{-1}$, $\pi_{31} = 100 \text{ min}^{-1}$, $\delta_1 = \delta_2 = \delta_3 = 0.01 \text{ min}^{-1}$, $p_{T1} = 1 \text{ nM}$, $p_{T2} = 1 \text{ nM}$, $p_{T3} = 1 \text{ nM}$. The dissociation constants for the top figure are $k_{d31} = 10 \text{ nM}$, $k_{d32} = 5 \text{ nM}$, and for the bottom figure are $k_{d31} = 1000 \text{ nM}$, $k_{d32} = 100 \text{ nM}$. The chemical reactions and the system models are given in Appendix D.3.

systems (with one slow variable and one fast variable) modeled by the CME, and points to two main sources of inaccuracy for the stochastic models obtained through standard QSSA [47]. The first source of inaccuracy arises from treating species that contain both slow and fast variables as purely slow variables. This is also a source of
Figure 6-4: Sample covariances computed using the Hill function based LNA model and the stochastic tQSSA+ based LNA model. Sample variances are computed using 150,000 simulation runs. The parameter values are \( \pi_{10} = 0.01 \text{ min}^{-1} \), \( \pi_{21} = 1 \text{ min}^{-1} \), \( \pi_{31} = 100 \text{ min}^{-1} \), \( \delta_1 = \delta_2 = \delta_3 = 0.01 \text{ min}^{-1} \), \( p_{T_1} = 1 \text{ nM} \), \( p_{T_2} = 1 \text{ nM} \), \( p_{T_3} = 1 \text{ nM} \). The dissociation constants for the top figure are \( k_{d31} = 10 \text{ nM} \), \( k_{d32} = 5 \text{ nM} \), and for the bottom figure are \( k_{d31} = 1000 \text{ nM} \), \( k_{d32} = 100 \text{ nM} \). The chemical reactions and the system models are given in Appendix D.3.

inaccuracy in the deterministic QSSA as noted by several studies [4, 15]. The second source of inaccuracy is disregarding the noise of the fast variable when taking the QSSA. In line with this, J.K. Kim et al. analyzed the steady state distribution and the variance of the fast variable and determine that the stochastic QSSA becomes
accurate when the sensitivity of the quasi-steady state expression to changes of the slow variable is small. This sensitivity term is equivalent to the retroactivity terms $R_{jl}$ in our model. Thus, our results recapitulate the findings of Kim’s steady state analysis on systems with two species, but, being based on general and rigorous derivation of the moment dynamics, our results are applicable to systems with more than two species and can also be used to study transient dynamics.

Additionally, a similar study has been performed by Thomas et al., where a reduced-order approximation for LNA is proposed[88]. They analyze several enzyme and gene network motifs and determine that the discrepancy between the quasi-steady state LNA models and the reduced-order LNA models proposed in their work arises mainly due to disregarding the noise of the fast variables. However, the reduced-order LNA developed in their work also regards some species concentrations that are combinations of fast and slow variables as purely slow variables. Therefore, their transient dynamics may not be accurately approximated by their reduction. In contrast to this, here, we identify conditions for which both the transient dynamics and the steady state are accurate.

Our analysis on the parameter conditions for validity of Hill function models have so far have only considered bounded values of protein concentrations and Hill functions. In future work, we aim to extend this analysis to capture a broader range of parameters and provide a more general condition for the validity of the Hill function models.

The reduced moment description obtained in this work can be used to accurately represent the moments of the original species dynamics even in the cases where the Hill function model may not be accurate. Thus, further analysis of this model may also lead to identification of stochastic properties of gene-regulatory networks that may not be captured with Hill function models.
6.5 Model derivations

6.5.1 Macroscopic reaction rate equations

As the starting point for writing LNA models is the deterministic reaction rate equations, we first derive the macroscopic reaction rates for a node $j$. Towards this end we first introduce the following sets: Let $\mathcal{Y}_{y_j}$ denote the set of nodes that take $y_j$ as an input. Let $\mathcal{B}_{ij}$ denote the set of all complexes in node $i$ that $y_j$ is bound.

Then, the macroscopic reaction rate equations for node $j$ can be written as

$$ y_j = \sum_{D_{jk} \in \mathcal{B}_j} \pi_{jk} d_{jk} - \delta_j y_j + \sum_{p \in \mathcal{Y}_{y_j}} \left( \sum_{D_{pk} \in \mathcal{Y}_{y_j} \setminus \mathcal{B}_{pj}} -\alpha_{pj} y_j^n d_{pk} + \sum_{D_{pj} \in \mathcal{B}_{pj}} \beta_{pj} d_{pj} \right), \quad (6.6) $$

$$ \dot{d}_{jl} = \sum_{Y_q \in \mathcal{B}_{zl}} \left( \sum_{\{D_{ik} \in \mathcal{Y}_{y_i} | \Theta_{ik} = \Theta_{zl} \}} \alpha_{jq} y_q^n d_{jk} - \beta_{jq} d_{jl} \right) $$

$$ - \sum_{Y_q \in \mathcal{B}_{zl}} \left( \alpha_{jq} y_q^n d_{jl} - \sum_{\{D_{jr} \in \mathcal{Y}_{y_j} | \Theta_{jl} = \Theta_{jr} \}} \beta_{jq} d_{jr} \right), \quad (6.7) $$

From (6.6), it follows that the dynamics for $y_j$ are dependent on the complexes of node $j$ and the complexes of nodes that take the TF $Y_j$ as an input. Therefore, we introduce

$$ c_{ij} = \sum_{D_{il} \in \mathcal{Y}_{y_i}} d_{il}, \quad \text{for } i \in \mathcal{Y}_{y_j} $$

where $c_{ij}$ is the sum of all complexes in node $i$ to which $y_j$ is bound. Furthermore, due to the conservation of the total amount of promoter in each node we have that, for a node $i$,

$$ p_{Ti} = c_{ij} + \sum_{D_{il} \in \mathcal{Y}_{y_i} \setminus \mathcal{B}_{ij}} d_{il}. $$

Then, we can write (6.6) in the form

$$ \dot{y}_j = \sum_{D_{jk} \in \mathcal{B}_j} \pi_{jk} d_{jk} - \delta_j y_j - \sum_{p \in \mathcal{Y}_{y_j}} (\alpha_{pj} y_j^n (p_{tp} - c_{pj}) - \beta_{pj} c_{pj}). \quad (6.8) $$

Then, from (6.7), we can see that the dynamics for each complex $d_{jk}$ are a function
of the other complexes in node $j$ and the input TFs to node $j$. Therefore let us denote by $d_i = [d_{i1}, \ldots, d_{i27}]$ the concentration vector of complexes in node $j$ and $u_j = [u_{j1}, \ldots, u_{jn}]$ be the concentration vector of input TFs to node $j$. Then, we can write the system (6.6) - (6.7) in the form

$$ y_j = f_j(y_j, d_j, t) - \sum_{p \in \mathcal{G}_j} g_{pj}(y_j, c_{pj}), \quad (6.9) $$

$$ \dot{d}_j = h_j(u_j, d_j), \quad (6.10) $$

where $f_j(y_j, d_j, t) = \sum_{Djk \in \mathcal{G}_j} \pi_{jk}d_{jk} - \delta_j y_j$, $g_{pj}(y_j, c_{pj}) = \alpha_{pj}y_j^n(p_{tp} - c_{pj}) - \beta_{pj}c_{pj}$, $h_j(u_j, d_j) = [h_{j1}(u_j, d_j); \ldots; h_{j27}(u_j, d_j)]$ where $h_{j1}(u_j, d_j) = \sum_{Y_{\epsilon} \in \Theta_{ji}} \left( \sum_{D_{ik} \in \mathcal{G}_j} \pi_{ik} \sum_{\Theta_{jt} \in \mathcal{G}_j} \alpha_{jq} y_j^n d_{jk} - \beta_{jq} d_{jt} \right)$ and

$$ \sum_{Y_{\epsilon} \in \Theta_{ji}} \left( \alpha_{jq} y_j^n d_{jt} - \sum_{D_{jr} \in \mathcal{G}_j} \beta_{jq} d_{jr} \right). $$

We further note that the complexes $c_{pj}$ represent the amount of $y_j$ bound in each node $p$. Therefore, the rate of change of $c_{pj}$ is given by

$$ \dot{c}_{pj} = g_{pj}(y_j, c_{pj}), $$

where $-g_{pj}(y_j, c_{pj})$ in (6.9) represents the rate of change of $y_j$, the free protein concentration, due to binding to node $p$.

Next, we use the dynamics for the deterministic trajectories derived in this section to obtain reduced-order LNA using the stochastic QSSA, which we call the Hill function model, and the reduced-order approximation for the moment dynamics using the stochastic tQSSA$^+$ introduced in Chapter 4. Towards this end we first present some definitions of two commonly occurring terms in these models.

6.5.2 Definitions

Let $\gamma_j(u_j) = [\gamma_{j1}(u_j), \ldots, \gamma_{j27}(u_j)]$ denote the solution to $\dot{d}_j = 0$, which represents the quasi-steady state expression for each promoter-TF complex. Then, the Hill
function is defined as

$$H_j(u_j) = \sum_{D_{jk} \in \mathcal{D}_j} \pi_{jk} \gamma_{jk}(u_j).$$  \hspace{1cm} (6.11)$$

Let $\omega_{pj}(y_j)$ denote the solution to $\dot{c}_{pj} = 0$, which represents the quasi-steady state expression for each total complex of $y_j$ in node $p$. Then, we define the retroactivity from node $p$ to node $j$ as

$$R_{jp}(y_j) = \frac{d\omega_{pj}(y_j)}{dy_j}.$$  

Using the expression of $g_{pj}(y_j, c_{pj})$ we find that

$$\omega_{pj}(y_j) = \frac{(p_{TI}/k_{di})}{(y_j^n / k_{di} + 1)},$$

where $k_{di} = \beta_{pj} / \alpha_{pj}$ is the dissociation constant between $y_j$ and the promoter binding site on node $p$. This leads to

$$R_{jp}(y_j) = \frac{(p_{TI} / k_{di}) y_j^{n-1}}{(y_j^n / k_{di} + 1)^2}.$$  \hspace{1cm} (6.12)$$

Next we present the LNA model based on stochastic QSSA.

6.5.3 Hill function model

In the QSSA approximation it is assumed that the dynamics of the complexes $d_j$ are much faster than the dynamics of the proteins. Hence, the concentrations of the complexes are approximated by their quasi-steady state values, obtained when $d_j = 0$. Since $c_{ij} = \sum_{D_{ui} \in \mathcal{D}_i} d_{ui}$, we have that $\dot{c}_{ij} = \sum_{D_{ui} \in \mathcal{D}_i} \dot{d}_{ui}$, leading to $g_{pj}(y_j, c_{pj}) = 0$, for all $p$ when $d_i = 0$. Thus, taking $\dot{d}_j = 0$ and $\dot{c}_{ij} = 0$ in (6.10) - (6.10) the reduced-order dynamics for the protein $y_j$ can be written as

$$\dot{\bar{y}}_j = H_j(\bar{u}_j) - \delta_j \bar{y}_j,$$  \hspace{1cm} (6.13)$$
where we use \( \tilde{y}_j, \tilde{u}_j \) denote the concentrations obtained using the QSSA, and the hill function \( H_j(\tilde{u}_j) \) is defined in (6.11).

The stochastic QSSA approach involves directly using the reduced dynamics (6.13) to derive the stochastic dynamics for the systems. Therefore, next, we proceed to derive the stochastic fluctuations in the LNA model. Denoting by \( \xi \) the fluctuations around \( \tilde{y}_j \). Then, using the definition of the LNA we have

\[
\dot{\xi}_j = \sum_{q \in \mathcal{S}_j} \frac{\partial H(\tilde{u}_j)}{\partial \tilde{y}_q} \xi_q - \delta_j \xi_j + \sqrt{H(\tilde{u}_j) \Gamma_{j1} - \sqrt{\delta_j \tilde{y}_j \Gamma_{j2}}}. \tag{6.14}
\]

Then, the variance and the covariance of the fluctuations (6.14) can be derived as

\[
\frac{dE[\dot{\xi}_j^2]}{dt} = 2 \sum_{q \in \mathcal{S}_j} \frac{\partial H(\tilde{u}_j)}{\partial \tilde{y}_q} E[\xi_q \dot{\xi}_j] - 2\delta_j E[\xi_j^2] + H_j(\tilde{u}_j) + \delta_j \tilde{y},
\]

\[
\frac{dE[\dot{\xi}_j \dot{\xi}_r]}{dt} = (-\delta_j - \delta_r) E[\xi_j \xi_r] + \sum_{q \in \mathcal{S}_j} \frac{\partial H(\tilde{u}_j)}{\partial \tilde{y}_q} E[\xi_q \dot{\xi}_j] + \sum_{q \in \mathcal{S}_r} \frac{\partial H(\tilde{u}_j)}{\partial \tilde{y}_q} E[\xi_q \dot{\xi}_r], \quad \text{for } j \neq r.
\]

Next, we use the model reduction for approach LNA introduced Chapter 4 to derive a reduced-order LNA and the corresponding moment dynamics.

### 6.5.4 Stochastic tQSSA\(^+\) model

**System in the singular perturbation form**

To derive a reduced-order model using stochastic tQSSA\(^+\), we first require the system dynamics to be written in standard singular perturbation form. In the derivation of the Hill function model, it is assumed that the concentrations \( y_j \) are the slow variables, when in fact their dynamics consists of both slow and fast components. Thus, in order to represent the dynamics in the singular perturbation form, we require the identification of slow and fast variables, for which the dynamics evolve on well
separated time-scales. This can be done by performing a coordinate change

\[ x_j = y_j + \sum_{p \in \mathcal{P}_{y_j}} \sum_{l \in \mathcal{A}_{y_j}} d_{jl}, \]

where \( x_j \) represents the total - free and bound - concentrations of protein \( y_j \), as frequently done in the singular perturbation approximations and tQSSA approaches. As the binding reactions are much faster than protein production and decay we can define the small parameter \( \epsilon = \delta/\beta \), by choosing a decay rate \( \delta \) and unbinding constant \( \beta \). Then we can write \( \beta_{ji} = \frac{1}{\epsilon} \delta_{ji} \) and \( \alpha_{ji} = \frac{1}{\epsilon} \delta_{ji} \), where \( k_{dji} = \beta_{ji}/\alpha_{ji} \) are the dissociation constants. With this, we can write the equations for a node \( j \) as

\[ \dot{y}_j = \sum_{D_{jk} \in \mathcal{Y}_j} \pi_{jk} d_{jk} - \delta_{j} y_j \]

\[ + \sum_{p \in \mathcal{X}_j} \left( \sum_{D_{pk} \in \mathcal{P}_p \setminus \mathcal{A}_p} - \frac{1}{\epsilon} \frac{\delta_{p}}{\beta k_{d_{pk}}} y_p^n d_{pk} + \sum_{D_{pk} \in \mathcal{A}_p} \frac{1}{\epsilon} \frac{\delta_{p}}{\beta} d_{pl} \right), \quad (6.15) \]

\[ \dot{d}_{jl} = \sum_{Y_e \in \mathcal{Y}_j} \left( \sum_{i \in \mathcal{A}_j} \frac{1}{\epsilon} \frac{\delta_{ij}}{\beta k_{d_{ij}}} y_i^n d_{ij} - \frac{1}{\epsilon} \frac{\delta_{ij}}{\beta} d_{ij} \right) \]

\[ - \sum_{Y_e \in \{Y_j \setminus \Theta_{ij}\}} \left( \frac{1}{\epsilon} \frac{\delta_{j}}{\beta k_{d_{ij}}} y_{j}^n d_{jl} - \sum_{D_{jr} \in \mathcal{Y}_j \setminus \Theta_{ij} \setminus Y_e} \frac{1}{\epsilon} \frac{\delta_{j}}{\beta} d_{jr} \right). \quad (6.16) \]

Next, in order to write the system into singular perturbation form, denote by \( \mathbf{c}_j \) the vector of complexes \( c_{ij} \) such that \( i \in \mathcal{Y}_j \). Then, we can write the system dynamics in the form

\[ \dot{x}_j = f_j(x_j - \sum_{p \in \mathcal{X}_j} c_{pj}, d_j, t), \quad (6.17) \]

\[ \epsilon \dot{d}_j = \hat{h}_j(u_{jT}, a_{u}, d_j), \quad (6.18) \]

\[ \epsilon \dot{c}_j = \hat{g}_j(x_j, c_j), \quad (6.19) \]

where \( \hat{f}_j(x_j, c_j, d_j, t) = f_j(x_j - \sum_{p \in \mathcal{X}_j} c_{pj}, d_j, t), \hat{g}_j(x_j, c_j) = g_{pj}(x_j - \sum_{p \in \mathcal{X}_j} c_{pj} \cdot c_{pj}), \) \( u_{jT} \) denotes the total (free and bound) concentrations of the inputs to \( Y_j \) and \( a_{uj} \)
denotes the vector of promoter-TF complexes that result from the binding between the inputs of node \( j \) (the elements \( u_{ji} \) in vector \( \mathbf{u}_j \)) to other nodes that are also regulated by inputs \( u_{ji} \), i.e., the elements of \( \mathbf{a}_{uj} \) are \( c_{kl} \) where \( k \in \mathbf{y}_{uj} \) and \( l \in \mathbf{a}_{j} \).

Then representing by \( \mathbf{x} = [x_1, \ldots, x_N] \) the vector of protein concentrations in the network and by \( \mathbf{d} = [d_{11}, \ldots, d_{12}m, d_{21}, \ldots, d_{22}m, \ldots] \), the vector of promoter-TF complexes we can write

\[
\begin{align*}
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{c}, \mathbf{d}, t) \\
\dot{\mathbf{d}} &= \mathbf{h}(\mathbf{x}, \mathbf{d}) \\
\dot{\mathbf{c}} &= \mathbf{g}(\mathbf{x}, \mathbf{c})
\end{align*}
\]

where

\[
\begin{align*}
\mathbf{f}(\mathbf{x}, \mathbf{c}, \mathbf{d}, t) &= [\mathbf{f}_1(x_1, c_1, d_1, t); \ldots; \mathbf{f}_j(x_N, c_N, d_N, t)] \\
\mathbf{h}(\mathbf{x}, \mathbf{d}) &= [\mathbf{h}_1(u_{1T}, a_{u1}, d_1); \ldots; \mathbf{h}_N(u_{TN}, a_{uN}, d_N)] \\
\mathbf{g}(\mathbf{x}, \mathbf{c}) &= [\mathbf{g}_1(x_1, c_1); \ldots; \mathbf{g}_N(x_N, c_N)]
\end{align*}
\]

**Approximation for the slow variables**

Let \( \mathbf{\lambda}(\mathbf{x}) \) and \( \mathbf{\rho}(\mathbf{x}) \) be the solutions to \( \hat{\mathbf{h}}(\mathbf{x}, \mathbf{d}) = 0 \) and \( \hat{\mathbf{g}}(\mathbf{x}, \mathbf{c}) = 0 \). Then, we have the reduced system

\[
\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{\lambda}(\mathbf{x}), \mathbf{\rho}(\mathbf{x}), t).
\]  

(6.20)

Note that the elements of \( \mathbf{\lambda}(\mathbf{x}) \) are the solutions to \( \hat{\mathbf{g}}_i(x_i, c_i) = 0 \), which only depends on \( x_i \). Therefore, we have that \( \mathbf{\lambda}(\mathbf{x}) = [\lambda_1(x_1); \ldots; \lambda_N(x_N)] \). Similarly, we have that elements of \( \mathbf{\rho}(\mathbf{x}) \) are the solutions to \( \hat{\mathbf{h}}_i(u_{iT}, a_{ui}, d_i) \), which is only dependent on \( u_{iT} \) and \( a_{ui} \). As \( a_{ui} \) are the complexes to which elements of \( \mathbf{u}_i \) are bound, we have that \( a_{ui} \) are also only dependent on \( u_{iT} \). Therefore, we can write \( \mathbf{\rho}(\mathbf{x}) = [\rho_1(u_{1T}); \ldots; \rho_N(u_{NT})] \). Then, we can write the reduced-order dynamics for
node $j$ as

$$\dot{x}_j = \sum_{D_{jk} \in G_j} \pi_{jk} \rho_{jk}(u_{jT}) - \delta_j(x_j - \sum_{p \in \mathcal{V}_{y_j}} \lambda_{jp}(x_j)), \quad (6.21)$$

where $\rho_{jk}(u_{jT})$ and $\lambda_{jp}(x_j)$ are elements of $\rho_j(u_{jT})$ and $\lambda_j(x_j)$, respectively (We drop the bars for notational simplicity). The stability of the slow manifold required for the model reduction is shown in Appendix D.1 where we show the stability corresponding to the fast variables $c_{ij}$'s, which then also implies the stability corresponding to the variables $d_{ik}$'s as $d_{ik} \geq 0$ and $c_{ij} = \sum_{D_{ik} \in G_i} d_{ik}$, for $i \in \mathcal{V}_{y_j}$.

According to Chapter 4, we can directly use the reduced-order dynamics for the slow variables to derive the corresponding stochastic fluctuations of the LNA. Therefore defining by $\chi_j$ the stochastic fluctuations that correspond to $x_j$ we have

$$\dot{\chi}_j = \sum_{D_{jk} \in G_j} \pi_{jk} \left( \sum_{q \in \Theta_{jk}} \frac{\partial \rho_{jk}(u_j)}{\partial x_q} \chi_q - \delta_j(1 - \sum_{p \in \mathcal{V}_{y_j}} \frac{\partial \lambda_{jp}(u_p)}{\partial x_j}) \chi_j \right) + \sqrt{\sum_{D_{jk} \in G_j} \pi_{jk} \rho_{jk}(u_j)} \Gamma_{xj1} - \sqrt{\delta_j(x_j - \sum_{p \in \mathcal{V}_{y_j}} \lambda_{jp}(x_j))} \Gamma_{xj2}. \quad (6.22)$$

Then, the moments of slow variable are given by

$$\frac{d\mathbb{E}[\chi_j^2]}{dt} = 2 \sum_{D_{jk} \in G_j} \pi_{jk} \left( \sum_{q \in \Theta_{jk}} \frac{\partial \rho_{jk}(u_j)}{\partial x_q} \mathbb{E}[\chi_q \chi_j] - 2 \delta_j(1 - \sum_{p \in \mathcal{V}_{y_j}} \frac{\partial \lambda_{jp}(u_p)}{\partial x_j}) \mathbb{E}[\chi_j^2] \right)$$

$$+ \sum_{D_{jk} \in G_j} \pi_{jk} \rho_{jk}(u_j) + \delta_j(x_j - \sum_{p \in \mathcal{V}_{y_j}} \lambda_{jp}(x_j)), \quad (6.23)$$

$$\frac{d\mathbb{E}[\chi_j \chi_r]}{dt} = \sum_{D_{jk} \in G_j} \pi_{jk} \left( \sum_{q \in \Theta_{jk}} \frac{\partial \rho_{jk}(u_j)}{\partial x_q} \mathbb{E}[\chi_q \chi_r] + \sum_{k=1}^{n} \pi_{rk} \sum_{q \in \Theta_{rk}} \frac{\partial \rho_{rk}(u_r)}{\partial x_q} \mathbb{E}[\chi_q \chi_j] \right)$$

$$- (\delta_r(1 - \sum_{p \in \mathcal{V}_{y_r}} \frac{\partial \lambda_{rp}(u_p)}{\partial x_r}) + \delta_j(1 - \sum_{p \in \mathcal{V}_{y_j}} \frac{\partial \lambda_{jp}(u_p)}{\partial x_j})) \mathbb{E}[\chi_r \chi_j] \quad (6.24)$$

The moments of the slow variables $\chi_j$ correspond to the moments of the total protein concentrations $x_j$. Thus, in order to derive the moments of $y_j = x_j - \sum_{p \in \mathcal{V}_{y_j}} c_{pj}$ we require an approximation for the variables $c_{ij}$ and their moments.
**Approximation for the fast variables**

We have that the dynamics of $c_{ij}$ for $i \in \mathcal{Y}_{ij}$ are given by

$$
\epsilon c_{ij} = \frac{\delta_j \beta_{ij}}{\beta k_{dij}} (x_j - \sum_{p \in \mathcal{Y}_{ij}} c_{pj})^n (p_{T_i} - c_{ij}) - \frac{\delta_j \beta_{ij}}{\beta} c_{ij}.
$$

(6.25)

Denote by $\zeta_{ij}$ the stochastic fluctuations corresponding to the complexes $c_{ij}$. Then, to derive the dynamics of the stochastic fluctuations $\zeta_{ij}$ we can directly use the definition of the LNA on the dynamics of $c_{ij}$ as the derivative is a linear operator and due to the fact that sum of normal random variables also result in a normal random variable where the variance is the sum of individual variance. Therefore, we have that

$$
\epsilon \dot{\zeta}_{ij} = n \frac{\delta_j \beta_{ij}}{\beta k_{dij}} (p_{T_i} - c_{ij}) (x_j - \sum_{p \in \mathcal{Y}_{ij}} c_{pj})^{n-1} \chi_j - (\delta_j \beta_{ij} \frac{\beta}{\beta k_{dij}} (x_j - \sum_{p \in \mathcal{Y}_{ij}} c_{pj})^n) \zeta_{ij}
$$

$$
- n \frac{\delta_j \beta_{ij} (p_{T_i} - c_{ij}) (x_j - \sum_{p \in \mathcal{Y}_{ij}} c_{pj})^{n-1}}{\beta k_{dij}} \sum_{p \in \mathcal{Y}_{ij}} \zeta_{pj}
$$

$$
+ \sqrt{\frac{\delta_j \beta_{ij} (x_j - \sum_{p \in \mathcal{Y}_{ij}} c_{pj})^n (p_{T_i} - c_{ij}) \Gamma_{c_{ij}}}{\beta k_{dij}}} - \sqrt{\frac{\delta_j \beta_{ij}}{\beta} c_{ij} \Gamma_{c_{ij}}}
$$

Next, we proceed to derive the reduced approximations of the fast variables $c_{ij}$ and $\zeta_{ij}$ when $\epsilon = 0$.

First considering the deterministic dynamics (6.25) and using $y_j = x_j - \sum_{p \in \mathcal{Y}_{ij}} c_{pj}$ we obtain

$$
\epsilon c_{ij} = \frac{\delta_j \beta_{ij}}{\beta k_{dij}} y_j^n (p_{T_i} - c_{ij}) - \frac{\delta_j \beta_{ij}}{\beta} c_{ij}
$$

(6.26)

and setting $\epsilon = 0$ we have

$$
c_{ij} = \frac{y_j^n p_{T_i}}{y_j^n + k_{dij}}.
$$

(6.27)

Next, we consider the stochastic fluctuations $\zeta_{ij}$. Note that for a given $\zeta_{ij}$ the dynamics are dependent only on the $\zeta_{pj}$ where $\zeta_{pj} \in \sum_{p \in \mathcal{Y}_{ij}}$. Therefore, let <
\( j_1, \ldots, j_{|\mathcal{V}_{y_j}|} \) be an ordered list of the node indexes of the proteins in \( \mathcal{V}_{y_j} \). Then, considering the vector \( \zeta_j = [\zeta_{j_1}; \ldots; \zeta_{j_{|\mathcal{V}_{y_j}|}}] \), the dynamics can be written as

\[
\begin{bmatrix}
\zeta_{j_1} \\
\vdots \\
\zeta_{j_{|\mathcal{V}_{y_j}|}}
\end{bmatrix} =
\begin{bmatrix}
n \frac{n \delta_j \beta_{j_1}}{\beta k_{d_{j_1}}} (p_{T_{j_1}} - c_{j_1}) y_j^{n-1} \\
\vdots \\
n \frac{n \delta_j \beta_{j_{|\mathcal{V}_{y_j}|}}}{\beta k_{d_{j_{|\mathcal{V}_{y_j}|}}}^{y_j}} (p_{T_{j_{|\mathcal{V}_{y_j}|}}} - c_{j_{|\mathcal{V}_{y_j}|}}) y_j^{n-1}
\end{bmatrix} \chi_j
\]

\[
- \begin{bmatrix}
(\delta_j \frac{\beta_{j_1}}{\beta} + \frac{\delta_j \beta_{j_1}}{\beta k_{d_{j_1}}} (x_j - \sum_{p \in \mathcal{V}_{y_j}} c_{pj}) y_j^{n-1}) + n \frac{\delta_j \beta_{j_1}}{\beta k_{d_{j_1}}} (p_{T_{j_1}} - c_{j_1}) y_j^{n-1} \\
\vdots \\
(\delta_j \frac{\beta_{j_{|\mathcal{V}_{y_j}|}}}{\beta} + \frac{\delta_j \beta_{j_{|\mathcal{V}_{y_j}|}}}{\beta k_{d_{j_{|\mathcal{V}_{y_j}|}}}^{y_j}} (p_{T_{j_{|\mathcal{V}_{y_j}|}}} - c_{j_{|\mathcal{V}_{y_j}|}}) y_j^{n-1} + n \frac{\delta_j \beta_{j_{|\mathcal{V}_{y_j}|}}}{\beta k_{d_{j_{|\mathcal{V}_{y_j}|}}}^{y_j}} (p_{T_{j_{|\mathcal{V}_{y_j}|}}} - c_{j_{|\mathcal{V}_{y_j}|}}) y_j^{n-1}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\zeta_{j_1} \\
\vdots \\
\zeta_{j_{|\mathcal{V}_{y_j}|}}
\end{bmatrix}
\]

\[
A(y_j) \chi_j + B(y_j) \zeta_j + \sqrt{\epsilon \sigma_1 (y_j) \Gamma_{c1}} + \sqrt{\epsilon \sigma_2 (y_j) \Gamma_{c2}}.
\]

From Theorem 4.1 Chapter 4, we can find approximation to the \( \zeta_j \) in the form

\( \zeta_j = \gamma_{2j}(y_j) \chi_j + \nu(y_j) N_j(0,1) \), given that the matrix \( B \) is Hurwitz. Note that, for computational simplicity, we choose to represent the functions \( \gamma_{2j} \) and \( \nu \) as functions of original variable \( y_j \) instead of the slow variables \( x_j \).

To find \( \gamma_{2j}(y_j) \) we set \( \epsilon = 0 \) which yields

\[
0 =
\begin{bmatrix}
n \frac{(p_{T_{j_1}}/k_{d_{j_1}}) y_j^{n-1}}{(y_j/k_{d_{j_1}})^{y_j+1}} \\
\vdots \\
(n \frac{(p_{T_{j_{|\mathcal{V}_{y_j}|}}/k_{d_{j_{|\mathcal{V}_{y_j}|}}}^{y_j}) y_j^{n-1}}{(y_j/k_{d_{j_{|\mathcal{V}_{y_j}|}}^{y_j}})^{y_j+1}} \\
(n \frac{\gamma_{j_1}}{\beta} y_j^{n-1}) \\
\end{bmatrix} \chi_j
\]
Noting that $R_{ji}(y_j) = n \frac{(p_{Tj_i} / k_{dj_i}) y_j^{n-1}}{(y_j^n / k_{dj_j}^{i+1})}$ as introduced in (6.12) we can write

$$0 = \begin{bmatrix}
    n \frac{(p_{Tj_i} / k_{dj_i}) y_j^{n-1}}{(y_j^n / k_{dj_j}^{i+1})} & & & & \\
    & \vdots & & & \\
    & & n \frac{(p_{Tj_i} / k_{dj_i}) y_j^{n-1}}{(y_j^n / k_{dj_j}^{i+1})} & & \\
    & & & \vdots & \\
    & & & & 1 + R_{ji}(y_j) / R_{ji}(y_j)
\end{bmatrix} \chi_j - \begin{bmatrix}
    n \frac{(p_{Tj_i} / k_{dj_i}) y_j^{n-1}}{(y_j^n / k_{dj_j}^{i+1})} & & & & \\
    & \vdots & & & \\
    & & n \frac{(p_{Tj_i} / k_{dj_i}) y_j^{n-1}}{(y_j^n / k_{dj_j}^{i+1})} & & \\
    & & & \vdots & \\
    & & & & 1 + R_{ji}(y_j) / R_{ji}(y_j)
\end{bmatrix} ^{-1} \begin{bmatrix}
    \zeta_{jj_1} \\
    \vdots \\
    \zeta_{jj_n}
\end{bmatrix}.$$

As the matrix $B$ is Hurwitz (Appendix D.1), we have that there exists a unique solution to (6.28) in the form $\zeta_j = B^{-1} A \chi_j = \hat{B}^{-1} D^{-1} A \chi_j$. Then, using the Claim
D.1 in Appendix D, with $a_i = \frac{1 + R_{ii}(y_i)}{R_{ij}(a_i)}$ we have

$$\hat{B}^{-1} = \begin{bmatrix}
R_{jj_1}(y_j) & 0 & \ldots & 0 \\
0 & R_{jj_2}(y_j) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & R_{jj_{|\mathcal{N}_y|}}(y_j)
\end{bmatrix} - \left( 1 + \sum_{i=1}^{|\mathcal{N}_y|} R_{jj_i}(y_j) \right)^{-1} \begin{bmatrix}
R_{jj_1}(y_j) & R_{jj_1}(y_j) & R_{jj_2}(y_j) & \ldots & R_{jj_1}(y_j) & R_{jj_{|\mathcal{N}_y|}}(y_j) \\
R_{jj_2}(y_j) & R_{jj_2}(y_j) & R_{jj_2}(y_j) & \ldots & R_{jj_2}(y_j) & R_{jj_{|\mathcal{N}_y|}}(y_j) \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
R_{jj_{|\mathcal{N}_y|}}(y_j) & R_{jj_{|\mathcal{N}_y|}}(y_j) & R_{jj_{|\mathcal{N}_y|}}(y_j) & \ldots & R_{jj_{|\mathcal{N}_y|}}(y_j) & R_{jj_{|\mathcal{N}_y|}}(y_j)
\end{bmatrix}.$$

(6.29)

Noting that $D^{-1}A$ is of the form

$$D^{-1}A = \begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix},$$

we have that

$$\gamma_{jj_i}(y_j) = (R_{jj_i}(y_j)) - \left( 1 + \sum_{i=1}^{|\mathcal{N}_y|} R_{jj_i}(y_j) \right)^{-1} \sum_{i=1}^{|\mathcal{N}_y|} R_{jj_i}(y_j) R_{jj_i}(y_j) \chi_j,$$

$$= \frac{R_{jj_i}(y_j)}{1 + \sum_{i=1}^{|\mathcal{N}_y|} R_{jj_i}(y_j) \chi_j} \chi_j.$$

Next to find an approximation for $\nu_j(y_j)$ we consider the Lyapunov equation

$$B(y_j)\nu_j(y_j)\nu_j(y_j)^T + \nu_j(y_j)\nu_j(y_j)^T B(y_j) = -(\sigma_1(y_j) + \sigma_2(y_j)),$$

(6.30)
where we use the fact that the $\Gamma_{c1}$ and $\Gamma_{c2}$ are independent Gaussian white noise processes.

Note that using the expression for $c_{ij}$ in (6.27) and simplifying further, the equation (6.30) can be written in the form

$$D(y_i)\dot{B}(y_j)\nu_j(y_j)\nu_j(y_j)^T + \nu_j(y_j)\nu_j(y_j)^TD(y_i)\dot{B}(y_j) = 2D(y_j), \quad (6.31)$$

Since the matrix $B(y_j)$ is Hurwitz (Appendix D.1) there exists a unique solution to (6.30) and thus to (6.31). Therefore, by substitution, we have that the unique solution to is in the form $\nu_j(y_j)\nu_j(y_j)^T = B(y_j)^{-1}$, where $B(y_j)^{-1}$ was found in (6.29).

Then, the elements of $\nu_j(y_j)$ are given by

$$\nu_{ji}(y_j) = \left( R_{ji}(y_j) - \frac{1}{1 + \sum_{l=1}^{[\mathbb{y}_j]} R_{ji}(y_j)^2} R_{ji}(y_j)^2 \right) y_i,$$

and for $l \neq i$,

$$\nu_{ji}(y_j) = -\frac{R_{ji}(y_j) R_{jj}(y_j)}{1 + \sum_{l=1}^{[\mathbb{y}_j]} R_{ji}(y_j)^2} y_j = -\frac{R_{ji}(y_j) R_{jj}(y_j)}{1 + \sum_{l=\mathbb{y}_j} R_{ji}(y_j)^2} y_j.$$

Next, we use the approximations derived for $c_{ij}$ and $\zeta_{ij}$ to obtain the moment dynamics of the variables $y_j$ and $\chi_i$.

**Approximation for the mixed variables**

Using the change of variables $y_j = x_j - \sum_{p \in \mathbb{y}_j} c_{pj}$, we can write

$$\dot{y}_j = \dot{x}_j - \sum_{p \in \mathbb{y}_j} c_{pj},$$

$$= (1 - \sum_{p \in \mathbb{y}_j} \frac{d\lambda_{pj}(x_j)}{dx_j}) \frac{dx_j}{dt},$$

120
From the definitions of $\lambda_{jp}(x_j)$ and $\omega_{jp}(y_j)$ we have that $\lambda_{jp}(x_j) = \omega_{jp}(y_j)$, and we can write

$$\frac{d\lambda_{jp}(x_j)}{dx_j} = \frac{d\lambda_{jp}(x_j)}{dy_j} \frac{dy_j}{dx_j},$$

and using $y_j = x_j - \sum_{y_p \in \mathcal{Y}_j} \omega_{pj}$ we have $\frac{dy_j}{dx_j} = \frac{1}{\sum_{y_p \in \mathcal{Y}_j} \frac{d\omega_{pj}}{dy_j}}$. Therefore, we have

$$y_j = \dot{x}_j - \sum_{y_p \in \mathcal{Y}_j} c_{pj},$$

$$= \frac{1}{1 + \sum_{y_p \in \mathcal{Y}_j} R_{jp}(y_j)} \left( \sum_{D_{jk} \in \mathcal{D}_j} \pi_{jk} \gamma_{jk}(u_j) - \delta_j y_j \right),$$

where we use the fact that $R_{jp}(y_j) = \omega_{jp}(y_j)$ denotes the retroactivity as defined in (6.12) and $\rho_{jk}(u_{jT}) = \gamma_{jk}(u_j)$.

Next considering the original variables $\xi_j = x_j - \sum_{p \in \mathcal{Y}_j} \zeta_{pj}$, we have

$$\mathbb{E}[\xi_j^2] = \mathbb{E} \left[ (x_j - \sum_{p \in \mathcal{Y}_j} \zeta_{pj})^2 \right],$$

$$= \mathbb{E}[\chi_j^2] - 2 \sum_{p \in \mathcal{Y}_j} \mathbb{E}[\chi_j \zeta_{pj}] + \sum_{p \in \mathcal{Y}_j} \mathbb{E}[\zeta_{pj}^2] + \sum_{p \in \mathcal{Y}_j} \sum_{r \in \mathcal{Y}_j \backslash p} \mathbb{E}[(\zeta_{pq} \zeta_{rj})].$$

Then, using the fast variable approximations for $\zeta_{pj}$ we have

$$\mathbb{E} \left[ \zeta_{pj}^2 \right] = \left( \frac{1}{1 + \sum_{l \in \mathcal{Y}_j} R_{jl}(y_j)} \right)^2 \mathbb{E}[\chi_j^2] + \sum_{l \in \mathcal{Y}_j} \frac{R_{jl}(y_j) + \sum_{i \in \mathcal{Y}_j \backslash l} R_{ji}(y_i) R_{jl}(y_j)}{1 + \sum_{l \in \mathcal{Y}_j} R_{jl}(y_j)} y_j$$

$$+ \sum_{i \in \mathcal{Y}_j} \sum_{k \in \mathcal{Y}_j \backslash i} \frac{R_{ji}(y_j) R_{jk}(y_j)}{1 + \sum_{l \in \mathcal{Y}_j} R_{jl}(y_j)} y_j,$$

$$= \left( \frac{1}{1 + \sum_{l \in \mathcal{Y}_j} R_{jl}(y_j)} \right)^2 \mathbb{E}[\chi_j^2] + \sum_{i \in \mathcal{Y}_j} \frac{R_{ji}(y_i)}{1 + \sum_{l \in \mathcal{Y}_j} R_{jl}(y_j)} y_j,$$

Similarly, we have

$$\mathbb{E}[\xi_j \xi_k] = \mathbb{E} \left[ (x_j - \sum_{p \in \mathcal{Y}_j} \zeta_{pj})(x_k - \sum_{r \in \mathcal{Y}_k} \zeta_{rk}) \right].$$
Then, taking the time derivative we obtain

\[
\frac{d\mathbb{E}[\xi_j^2]}{dt} = \frac{d}{dt} \left( \frac{1}{1 + \sum_{l \in \mathcal{y}_j} R_j(y_j)} \right)^2 \mathbb{E}[\chi_j^2] + \frac{1}{1 + \sum_{l \in \mathcal{y}_j} R_j(y_j)} \frac{d}{dt} \sum_{l \in \mathcal{y}_j} R_j(y_j) \mathbb{E}[\chi_j^2] + \frac{1}{1 + \sum_{l \in \mathcal{y}_j} R_j(y_j)} \sum_{l \in \mathcal{y}_j} \frac{dR_j(y_j)}{dy_j} \frac{dy_j}{dt} \mathbb{E}[\chi_j^2]
\]

\[
+ \frac{1}{1 + \sum_{l \in \mathcal{y}_j} R_j(y_j)} \frac{d}{dt} \frac{1}{1 + \sum_{l \in \mathcal{y}_j} R_j(y_j)} \mathbb{E}[\chi_j^2] + \frac{1}{1 + \sum_{l \in \mathcal{y}_j} R_j(y_j)} \sum_{l \in \mathcal{y}_j} \frac{dR_j(y_j)}{dy_j} \frac{dy_j}{dt} \mathbb{E}[\chi_j^2] + \frac{1}{1 + \sum_{l \in \mathcal{y}_j} R_j(y_j)} \sum_{l \in \mathcal{y}_j} \frac{dR_j(y_j)}{dy_j} \frac{dy_j}{dt} \mathbb{E}[\chi_j^2]
\]

Then, substituting the moment dynamics for the slow variables from (6.23)-(6.24), we obtain the moment dynamics

\[
\frac{d\mathbb{E}[\xi_j^2]}{dt} = \frac{1}{1 + R_j(y_j)} \left[ -2\delta_j(1 + F_j(y_j))\mathbb{E}[\xi_j^2] + 2\sum_{q \in \mathcal{x}_j} \frac{\partial H_j(u_j)}{\partial x_q} \mathbb{E}[\xi_q \xi_j] 
+ H_j(u_j)(1 + G_j(y_j)) + \delta_j y_j(1 - G_j(y_j)) - J_j(y_j) \right],
\]

(6.32)

\[
\frac{d\mathbb{E}[\xi_j \xi_r]}{dt} = \frac{1}{1 + R_j(y_j)} \sum_{q \in \mathcal{x}_j} \frac{\partial H_j(u_j)}{\partial y_q} \mathbb{E}[\xi_q \xi_r] + \frac{1}{1 + R_r(y_r)} \sum_{q \in \mathcal{x}_r} \frac{\partial H_r(u_r)}{\partial y_q} \mathbb{E}[\xi_q \xi_r] 
- \left( \delta_r \frac{1}{1 + R_r(y_r)} (1 - F_r(y_r)) + \delta_j \frac{1}{1 + R_j(y_j)} (1 - F_j(y_j)) \right) \mathbb{E}[\xi_j \xi_r] - I_{jr}(y_j, y_r),
\]
where

\[ R_{ji}(y_j) = \frac{e^{(p_{ri}/k_i)d_j}y_j^{n-1}}{(y_j^n/k_i + 1)^2}, \]

\[ \mathcal{R}_j(y_j) = \sum_{i \in \mathcal{Y}_j} R_{ji}(y_j), \]

\[ F_j(y_j) = \frac{1}{\delta_j} \sum_{i \in \mathcal{Y}_j} \frac{dR_{ji}(y_j)}{dy_j} \frac{dy_j}{dt}, \]

\[ G_j(y_j) = \frac{(2 \sum_{i \in \mathcal{Y}_j} R_{ji}(y_i) + 1) \sum_{i \in \mathcal{Y}_j} \frac{dR_{ji}(y_i)}{dy_j}}{(1 + \sum_{i \in \mathcal{Y}_j} R_{ji}(y_j))^2} y_j, \]

\[ J_j(y_j) = 2 \frac{\partial H_j(u_j)}{\partial y_j} \frac{\sum_{i \in \mathcal{Y}_j} R_{ji}(y_i)}{1 + \sum_{i \in \mathcal{Y}_j} R_{ji}(y_j)} y_j, \]

\[ I_{jr}(y_j, y_{jr}) = \frac{1}{1 + \sum_{i \in \mathcal{Y}_j} R_{ji}(y_j)} \frac{\partial H_j(u_j)}{\partial y_r} \frac{\sum_{i \in \mathcal{Y}_r} R_{ri}(y_i)}{1 + \sum_{i \in \mathcal{Y}_r} R_{ri}(y_{jr})} y_{jr} \]
\[ + \frac{1}{1 + \sum_{i \in \mathcal{Y}_r} R_{ri}(y_r)} \frac{\partial H_r(u_r)}{\partial y_j} \frac{\sum_{i \in \mathcal{Y}_j} R_{ji}(y_j)}{1 + \sum_{i \in \mathcal{Y}_j} R_{ji}(y_{jr})} y_{jr}. \]
Chapter 7

Conclusion

7.1 Contributions

In this thesis, we considered the problem of model order reduction for SDE representations of biomolecular systems: the chemical Langevin equation and the Linear Noise Approximation. Under time-scale separation conditions we obtained a reduced-order CLE for systems with linear propensity functions and a reduced-order LNA which can be used to model systems with both linear and nonlinear propensity functions. The main contributions of our results are as follows:

- *Error quantification between full and reduced systems*
  
The existing methods for model reduction CLE and LNA based on time-scale separation do not provide an error quantification between the full and the reduced systems. In contrast, in this work, we use the moment dynamics to provide an error quantification between the full system dynamics and the reduced-order systems. In particular, we show that the moments of the reduced-order model are in an $O(\epsilon)$-neighborhood of the moments of the full system, where $\epsilon$ is the singular perturbation parameter. For the case of chemical Langevin equation, these results hold for moments of all orders of the slow variable and the first and second moments of the fast variable. For the case of LNA, we prove convergence for the first and second moments, and since the stochastic fluctu-
ations considered in the LNA are normal random variables, convergence of the first and second moments imply convergence of all moments and convergence in distribution.

We note that the convergence of moments is in general a weaker condition than convergence in mean or probability as considered by other studies of stochastic singular perturbation that do not fit the form of CLE or the LNA. However, stochastic properties of biomolecular systems are often studied using the mean and variance of chemical species. Thus, it suffices to consider the convergence of moments to give a good approximation of the statistical properties of the full system.

- **Slow and fast variable approximations**
  We provide approximations for both slow and fast variables, as opposed to existing methods in the literature that only provide approximations for the slow variables. In biomolecular systems, species dynamics often consists of both slow and fast components. Thus, it is crucial to obtain an approximation for the fast variables in order to utilize the reduce-order model to study species with such ‘mixed’ dynamics.

- **Easy computation of the reduced systems**
  The reduced-order models developed in our work can be computed by solving a set of algebraic equations similar to singular perturbation techniques in deterministic systems. The stability assumptions that we require are also equivalent to the conditions that would be encountered in the application of singular perturbation to a deterministic system model.

- **Validity of stochastic tQSSA method in the literature**
  Our results on the reduced-order model for LNA also provide a rigorous justification for the validity of stochastic total quasi-steady state in the literature (tQSSA). We show that the reduced order LNA model obtained through stochastic tQSSA is equivalent to the slow variable approximation obtained in
our work. This implies that the reduced-order model obtained using stochastic tQSSA also provides a good approximation for the slow variables of full system. Furthermore, this indicates that when the molecular counts are sufficiently large, the stochastic tQSSA yield a valid approximation for CME under time-scale separation, similar to the observations in previous studies [47, 48].

We then considered two applications of the model reduction techniques developed in this work.

In the first application, we analyzed the interplay between modularity and signal noise in biomolecular network. Modularity allows the interconnection biomolecular components without causing a perturbation in the signal transmitted between the components. However, conditions for modularity often imply higher noise in the signals. We considered several network motifs that commonly occur in both natural and synthetic systems and used reduced-order CLE models and LNA models to quantify the stochastic effects of these systems under conditions that guarantee modularity of the network components. We find that while there are often trade-offs between high modularity and signal noise, certain methods of ensuring modularity such as the use of high-gain negative feedback can attenuate the signal noise in downstream components. These findings together with the mathematical quantifications of the trade-offs can be used to inform the design of biological systems in synthetic biology.

In the second application, we considered the problem of obtaining reduced-order stochastic models for gene-regulatory networks. In particular, we use the reduced-order LNA developed in this work to obtain a set of reduced moment dynamics for gene-regulatory networks. We use these moments dynamics to investigate the validity of stochastic QSSA models that are used in the literature.

7.2 Future directions

The results on the model order reduction methods can be extended in several ways:

- **Extension to the chemical Master equation**

  An immediate extension of this research is the development of similar model
reduction techniques for the chemical Master equation. However, the derivation of the moment dynamics for a general biomolecular system with nonlinear propensity functions is complicated due to the fact that the system of moment equations are not closed. Although, there are many existing moment closure methods in the literature, these methods do not provide a measure for the goodness of the resulting set of closed moments. This leads to the need for robust moment closure methods as in the work of Naghnaeian and Del Vecchio [62].

Another possible direction involves the extension to other approximations of the CME, such as the mass fluctuation kinetic (MFK) equations developed in [25], which provide an approximate description for the mean and the variance of species in the system.

- **Tighter convergence criteria**

  In this work, we consider the convergence of moments between the full and the reduced systems, which in general do not imply other forms of stochastic convergence such as the convergence in mean or the convergence in probability. One way that these results could be extended is by considering specific systems for which convergence of moments imply other forms of stochastic convergence. For example, Prokhorov’s method of moments state that the convergence of a sequence of moments imply convergence of distribution if the limiting sequence of moments uniquely determine the corresponding random variable’s probability distribution. There are also studies on conditions for the uniqueness of the set of moments for random variables described by stochastic differential equations [82]. Such tools from the mathematical literature could be used to investigate the possible extensions of the convergence results.

  The investigation of convergence of autocorrelation function which takes the form $R(t, t') = \mathbb{E}[X(t)X(t')]$ for a stochastic process $X(t)$ could also provide possible extensions of convergence measures since the convergence of distribution of a sequence of stochastic process $X_n(t)$ to a stochastic process $X(t)$ is defined by the convergence of the cumulative distribution functions of the form
Another direction that the convergence properties could be tightened is by extending the results to infinite time-interval. The results presented in this work are valid for a finite time interval as we consider the application of Tikhonov's theorem on the finite time-interval. Further stability conditions on the reduced-order models could be used to extend these results to the infinite time-intervals.
Appendix A

Proofs pertaining to Chapter 3

A.1 Proof of Claim 3.1

In order to derive the dynamics of $E[x^{(k)}]$, consider the drift and diffusion terms of the slow variable dynamics of the full system (3.2)-(3.2), which can be denoted as $f_x(x, z, t) = [f_{x_1}(x, z, t), \ldots, f_{x_n}(x, z, t)]^T$ and $\sigma_x(x, z, t) = [\sigma_{ij}^x(x, z, t), t]$ for $i = 1, \ldots, n$ and $j = 1, \ldots, d_x$. Then, using the Ito formula as in [78, p. 86] the moment dynamics of $x^{(k)}$ for each $k = (k_1, \ldots, k_n) \in \mathcal{K}_{n_2}^P$ can be derived as

$$
\frac{dE[x^{(k)}]}{dt} = \sum_{i=1}^{n} k_i E[f_x(x, z, t) x_1^{k_1} \ldots x_i^{k_i-1} \ldots x_n^{k_n}]
+ \frac{1}{2} \sum_{p=1}^{n} k_p (k_p - 1) E[\phi_{pp}(x, z, t) x_1^{k_1} \ldots x_p^{k_p-2} \ldots x_n^{k_n}]
+ \sum_{l=2}^{n} \sum_{j=1}^{l-1} k_l k_j E[\phi_{lj}(x, z, t) x_1^{k_1} \ldots x_j^{k_j-1} \ldots x_l^{k_l-1} \ldots x_n^{k_n}],
$$

(A.1)

where $\phi_{ij}(x, z, t)$ for $i, j = 1, \ldots, n$ are the elements of the matrix $\Phi(x, z, t)$ defined in Assumption 3.2. From Assumptions 3.1 - 3.2, we have that the functions $f_x(x, z, t)$ and $\phi_{ij}(x, z, t)$ are affine in $x$ and $z$. Hence, it follows that the dynamics of $P^\text{th}$ order moments will depend only on moments of order up to $P$. Under Assumption 3.1, we also have that $A_3(t)$ and $B_3(t)$ are continuous functions. Therefore, for appropriate continuous functions $C_{1i}(t)$ and $C_{2ji}(t)$ for $i \in \mathcal{G}^P_{n_2}$, $l \in \mathcal{G}^1_{n_2}$ and $j \in \mathcal{G}^{P-1}_{n_2}$, the moment
dynamics in (A.1) can be written in the form given in (3.11).

In order to derive the dynamics of \( E[z(g)] \), we consider the drift and diffusion terms of the fast variable dynamics of the full system (3.2)-(3.2), which can be denoted as 
\[
(1/\epsilon)f_z(x, z, t) = (1/\epsilon)[f_{z_1}(x, z, t, \epsilon), \ldots, f_{z_n}(x, z, t, \epsilon)]^T
\]
and 
\[
(1/\sqrt{\epsilon})\sigma_z(x, z, t, \epsilon) = (1/\sqrt{\epsilon})[\sigma_z^{ij}(x, z, t, \epsilon)]
\]
for \( i = 1, \ldots, n \) and \( j = 1, \ldots, (d_x + d_f) \). Then, as in [78, p. 86] the moment dynamics of \( z^{(g)} \) for each \( g = (g_1, \ldots, g_{nf}) \in \mathcal{K}_{n_f}^P \) can be written as

\[
\frac{dE[z^{(g)}]}{dt} = \sum_{i=1}^{m} g_i E \left[ \frac{1}{\epsilon} f_z(x, z, t, \epsilon) z_i^{g_1} \ldots z_i^{g_{i-1}} z_i^{g_i} \right] + \frac{1}{2} \sum_{p=1}^{m} g_p (g_p - 1) E \left[ \frac{1}{\epsilon} \lambda_{pp}(x, z, t, \epsilon) z_i^{g_1} \ldots z_i^{g_{p-2}} z_i^{g_p} \right] + \sum_{l=2}^{m} \sum_{p=1}^{l-1} g_l g_p E \left[ \frac{1}{\epsilon} \lambda_{pl}(x, z, t, \epsilon) z_i^{g_1} \ldots z_i^{g_{l-2}} z_i^{g_l} \right],
\]

where \( \lambda_{ij}(x, z, t, \epsilon) \) for \( i, j = 1, \ldots, m \) are the elements of the matrix \( \Lambda(x, z, t, \epsilon) \) defined in Assumption 3.2. Under Assumptions 3.1–3.2, we have that the functions \( f_z(x, z, t, \epsilon) \) and \( \lambda_{ij}(x, z, t, \epsilon) \) are affine in \( x \) and \( z \) and are continuously differentiable in their arguments. Thus, it follows that the dynamics of the moments of order \( P \) will depend only on moments of order less than or equal to \( P \). Then, multiplying both sides of the equation (A.2) by \( \epsilon \), we can represent the moments \( E[z^{(g)}] \) for each \( g \in \mathcal{K}_{n_f}^P \) as in (3.9) for appropriate continuously differentiable functions \( D_{1a}(t) \) and \( D_{2bc}(t) \) for \( a \in \mathcal{G}_{n_f}^P, c \in \mathcal{G}_{n_f}^1, b \in \mathcal{G}_{n_f}^{P-1} \).

In order to find the dynamics of \( E[z^{(k_x)} x^{(k_x)}] \), we consider the vector 
\[
[z_1, \ldots, z_{nf}, x_1, \ldots, x_{nx}]^T, \text{ for which the drift and the diffusion terms can be represented by the vector } [f_{z_1}(x, z, t, \epsilon)/\epsilon, \ldots, f_{z_{nf}}(x, z, t, \epsilon)/\epsilon, f_{x_1}(x, z, t) \ldots f_{x_{nx}}(x, z, t)]^T \text{ and matrix } [(1/\sqrt{\epsilon})\sigma_z^{ij}(x, z, t, \epsilon); [\sigma_x(x, z, t) 0]^k_j] \text{ respectively, for } i = 1, \ldots, m, j = 1, \ldots, (d_x + d_f), k = 1, \ldots, n, \text{ in which } [\sigma_x(x, z, t) 0] \text{ denotes a matrix-valued function } \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{n \times (d_x + d_f)} \text{ and } [\sigma_z(x, z, t) 0]^k_j \text{ denotes the elements of the matrix. Then, from [78, p. 86], the moment dynamics of } z^{(k_x)} x^{(k_x)} \text{ for } k_x = (c_1, \ldots, c_{nf}) \in \mathcal{K}_{n_f}^{P_z} \text{ and } k_x = (k_1, \ldots, k_{nf}) \in \mathcal{K}_{n_x}^{P_z} \text{ can be written as}
\[ \mathbb{E}[z^{(k_s)} x^{(k_s)}] \]
\[ \frac{dt}{dt} \]
\[ \sum_{i=1}^{m} c_i \mathbb{E} \left[ \frac{1}{\epsilon} f_{z_i}(x, z, t, \epsilon) z_{1}^{c_i} \ldots z_{i}^{c_{i-1}} \ldots z_{n_f}^{c_{i}} x_{1}^{k_{i1}} \ldots x_{n_s}^{k_{is}} \right] \]
\[ + \sum_{i=1}^{n} k_i \mathbb{E} \left[ f_{z_i}(x, z, t) z_{1}^{c_{i1}} \ldots z_{n_f}^{c_{i}} x_{1}^{k_{i1}} \ldots x_{n_s}^{k_{is}} \right] \]
\[ + \frac{1}{2} \sum_{i=1}^{m} c_i (c_i - 1) \left[ \lambda_{ij}(x, z, t, \epsilon) z_{1}^{c_{i1}} \ldots z_{i}^{c_{i-2}} \ldots z_{n_f}^{c_{i}} x_{1}^{k_{i1}} \ldots x_{n_s}^{k_{is}} \right] \]
\[ + \frac{1}{2} \sum_{i=1}^{n} k_i (k_i - 1) \left[ \phi_{ik}(x, z, t, \epsilon) z_{1}^{c_{i1}} \ldots z_{n_f}^{c_{i}} x_{1}^{k_{i1}} \ldots x_{i}^{k_{i-2}} \ldots x_{n_s}^{k_{is}} \right] \]
\[ + \sum_{i=2}^{m} \sum_{j=1}^{m} c_i c_j \mathbb{E} \left[ \frac{\lambda_{ij}(x, z, t, \epsilon)}{\epsilon} z_{1}^{c_{i1}} \ldots z_{j}^{c_{j-1}} \ldots z_{n_f}^{c_{i1}} x_{1}^{k_{i1}} \ldots x_{n_s}^{k_{is}} \right] \]
\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} k_i k_j \mathbb{E} \left[ \frac{\theta_{ij}(x, z, t, \epsilon)}{\sqrt{\epsilon}} z_{1}^{c_{i1}} \ldots z_{j}^{c_{j-1}} \ldots z_{n_f}^{c_{i1}} x_{1}^{k_{i1}} \ldots x_{n_s}^{k_{is}} \right] \]
\[ + \sum_{i=2}^{n} \sum_{j=1}^{n} k_i k_j \mathbb{E} \left[ \phi_{ij}(x, z, t) z_{1}^{c_{i1}} \ldots z_{n_f}^{c_{i1}} x_{1}^{k_{i1}} \ldots x_{j}^{k_{j-1}} \ldots x_{n_s}^{k_{is}} \right] \]  \( (A.3) \)

where \( \lambda_{ij}(x, z, t, \epsilon), \phi_{ik}(x, z, t) \) and \( \theta_{ik}(x, z, t, \epsilon) \) for \( i, j = 1, \ldots m, \ l, k = 1, \ldots n \) are the elements of the matrices \( \Lambda(x, z, t, \epsilon), \Phi(x, z, t) \) and \( \Theta(x, z, t, \epsilon) \) defined in Assumption 3.2, respectively. We have that the functions \( f_{z_i}(x, z, t, \epsilon), f_{x_i}(x, z, t), \lambda_{ij}(x, z, t, \epsilon), \phi_{ik}(x, z, t) \) and \( \theta_{ik}(x, z, t, \epsilon) \) are affine in \( x \) and \( z \) and are continuously differentiable in their arguments due to Assumption 3.1 - 3.2. Thus, for appropriate functions \( F_{2a}(t, \epsilon), F_{1u}(t, \epsilon), F_{3qrk_s}(t, \epsilon) \), for \( a \in \mathcal{G}_{n_f}^P, u \in \mathcal{G}_{n_s}^P, q = 2, \ldots, P, r = 1, \ldots, q, k \in \mathcal{G}_{n_f}^r, s \in \mathcal{G}_{n_s}^{r-s} \) the dynamics of \( \mathbb{E}[z^{(k_s)} x^{(k_s)}] \) can be written in the form of (3.10).
A.2 Proof of Claim 3.2

To find the moment dynamics $E[x^{(k)}]$, note that the reduced slow system (3.5) is obtained by taking $z = \gamma(x, t)$ in (3.2). Thus, the dynamics for $E[x^{(k)}]$ for all $k = (k_1, \ldots, k_n) \in \mathbb{K}_{n_j}$ can be obtained by following the proof of Claim 3.1 with $x = \bar{x}$ and $z = \gamma(\bar{x}, t)$, which yields equation (3.11).

Next, to derive the moment dynamics of the reduced fast system in (3.6), we take the expectation of equation (3.6), which yields

$$E[\bar{z}] = E[\gamma(\bar{x}, t) + g(\bar{x}, t)N]. \quad (A.4)$$

From the definition of the reduced fast system, we have that $N$ is a vector of standard normal random variables independent of the random vector $\bar{x}$. Thus, we have that $E[N] = 0$ and as the expectation operator is linear, we obtain $E[\bar{z}] = E[\gamma(\bar{x}, t)]$, which yields equation (3.12) for the case where $P_z = 1$.

Next, considering the second moment of the reduced fast system we have that

$$E[z_iz_j] = E[(\gamma_i(\bar{x}, t) + \sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)N_l)(\gamma_j(\bar{x}, t) + \sum_{l=1}^{d_x+d_f} g_{jl}(\bar{x}, t)N_l)].$$

Expanding further, we obtain

$$E[z_iz_j] = E[\gamma_i(\bar{x}, t)\gamma_j(\bar{x}, t)] + E \left[ \gamma_i(\bar{x}, t) \sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)N_l \right] + E \left[ \sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)N_l \gamma_j(\bar{x}, t) \right]$$

$$+ E \left[ \left( \sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)N_l \right) \left( \sum_{l=1}^{d_x+d_f} g_{jl}(\bar{x}, t)N_l \right) \right].$$

Since the elements of the vector $N$ are independent standard normal random variables, we have that $E[N_i] = 0$ for all $i$, and $E[N_iN_j] = 0$ for $i \neq j$ and $E[N_iN_j] = 1$ for $i = j$.

Therefore, we obtain $E[z_iz_j] = E[\gamma_i(\bar{x}, t)\gamma_j(\bar{x}, t)] + E \left[ \sum_{l=1}^{d_x+d_f} g_{il}(\bar{x}, t)g_{jl}(\bar{x}, t) \right]$, which can be written in the form of equation (3.12) for the case where $P_z = 2$.  

134
A.3 Proof of Claim 3.3

From Claim 3.1, we note that the fast variables that appear in the slow variable dynamics in (3.8) are of the form $E[z_i x^{(j)}]$ for $j = (k_1, \ldots, k_n) \in \mathcal{G}_{n}^{P-1}$ and $i = 1, \ldots, m$. Thus, we first consider setting $\epsilon = 0$ in the dynamics of $E[z_i x^{(j)}]$. The dynamics of $E[z_i x^{(j)}]$ can be obtained from the derivation of (3.10) in the proof of Claim 3.1, with $k_z = (c_1, \ldots, c_{n_f}) \in \mathcal{K}_{n_f}^{1}$ which gives $c_i = 1$ and $c_l = 0$ for $l \neq i$ for each $i = 1, \ldots, m$. Then, we have

$$
\epsilon \frac{dE[z_i x^{(j)}]}{dt} = E[f_z(x, z, t, \epsilon)x^{(j)}] + \sum_{l=1}^{n} k_l E[f_{z_l}(x, z, t)z_l x_1^{k_1} \ldots x_{l-1}^{k_{l-1}} \ldots x_{n_s}^{k_{n_s}}] + \sum_{p=1}^{n} k_p (k_p - 1) E[\phi_{pp}(x, z, t)z_p x_1^{k_1} \ldots x_{p-1}^{k_{p-1}} \ldots x_{n_s}^{k_{n_s}}] + \sqrt{\epsilon} \sum_{m=1}^{n_f} k_{m_f} E[\theta_{mm}(x, z, t, \epsilon)x_1^{k_1} \ldots x_{n_f-1}^{k_{n_f-1}} \ldots x_{n_s}^{k_{n_s}}] + \epsilon \sum_{l=2}^{n_f} \sum_{p=1}^{k_p} k_l k_p E[\phi_{lp}(x, z, t)z_l x_1^{k_1} \ldots x_{p-1}^{k_{p-1}} \ldots x_{l-1}^{k_{l-1}} \ldots x_{n_s}^{k_{n_s}}],
$$

where $\phi_{ik}(x, z, t)$ and $\theta_{il}(x, z, t, \epsilon)$ are the elements of the matrices $\Phi(x, z, t)$ and $\Theta(x, z, t, \epsilon)$ defined in Assumption 3.2, where we have that $\lim_{\epsilon \to 0} \theta_{ij}(x, z, t, 0) < \infty$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Thus, setting $\epsilon = 0$ in the dynamics of the vector $E[z_i x^{(j)}]$ yields

$$
E[f_z(x, z, t, 0)x^{(j)}] = 0.
$$

Under Assumption 3.2 - 3.3, there exist a unique solution to equation (A.6), which is given by $E[z x^{(j)}] = -B_z^{-1}(B_1 E[x^{(j)}] + B_3(t) E[x^{(j)}])$. Considering the expression for $\gamma(x, t)$ in equation (3.4), it follows that $E[\gamma(x, t)x^{(j)}] = E[-B_z^{-1}(B_1 x + B_3(t))x^{(j)}] = E[z x^{(j)}]$. Thus, it can be seen that setting $\epsilon = 0$ in (A.5), we obtain

$$
E[z_i x^{(j)}] = E[\gamma_i(x, t)x^{(j)}], \quad i \in \{1, \ldots, m\}, \; j = (k_1, \ldots, k_n) \in \mathcal{G}_{n_s}^{P-1}.
$$
Then, substituting (A.7) in (3.8) yields the set of equations (A.5) for the moments of the slow variable x.

Considering the equation (A.7) with $P = 1$, we also obtain that $\mathbb{E}[z_i] = \mathbb{E}[\gamma_i(x, t)]$, which results in the equation (3.14) for the case where $P_z = 1$.

Next, we consider setting $\epsilon = 0$ in the second order moment dynamics of $z$ in (3.9) given by the case where $g \in K^2_{\alpha_f}$. We denote these moments by $\mathbb{E}[z_iz_j]$ for $i, j = 1, \ldots, m$, for which, the dynamics can be obtained from the derivation of (3.9) in the proof of Claim 3.1, taking $g = (g_1, \ldots, g_{\alpha_f}) \in K^2_{\alpha_f}$ with $g_i = 1, g_j = 1$ and $g_l = 0$ for all $l \neq i, j$. Then, representing the second moments of $z$ in matrix form we have

$$
\frac{d\mathbb{E}[zz^T]}{dt} = \mathbb{E}[zf(x, z, t, \epsilon)^T] + \mathbb{E}[f_z(x, z, t, \epsilon)z^T] + \mathbb{E}[\sigma_z(x, z, t, \epsilon)\sigma_z(x, z, t, \epsilon)^T]. \quad (A.8)
$$

Then, setting $\epsilon = 0$ in the equation (A.8) together with Assumptions 3.1 - 3.2, yields

$$
\mathbb{E}[zx^T]B_1^T + \mathbb{E}[zx^T]B_2^T + \mathbb{E}[z]B_3(t)^T + B_1\mathbb{E}[zx^T] \\
+ B_2\mathbb{E}[zx^T] + B_3(t)\mathbb{E}[x^T] + \Lambda(\mathbb{E}[x], \mathbb{E}[z], t, 0) = 0. \quad (A.9)
$$

From equation (A.7), we can write $\mathbb{E}[z] = \mathbb{E}[\gamma(x, t)]$ and $\mathbb{E}[zx^T] = \mathbb{E}[\gamma(x, t)x^T]$ for the case where $P = 1$ and $P = 2$, which can then be used in (A.9) to obtain

$$
\mathbb{E}[zx^T]B_2^T + B_2\mathbb{E}[zx^T] = -\mathbb{E}[\gamma(x, t)x^T]B_1^T - \mathbb{E}[\gamma(x, t)]B_3(t)^T \\
- B_1\mathbb{E}[\gamma(x, t)x^T]^T - B_3(t)\mathbb{E}[\gamma(x, t)]^T - \Lambda(\mathbb{E}[x], \mathbb{E}[\gamma(x, t)], t, 0). \quad (A.10)
$$

The equation (B.14) is in the form of the Lyapunov equation

$$
A^TP + PA = -Q,
$$

136
with

\[ P = \mathbb{E}[zz^T], \]
\[ Q = -\mathbb{E}[\gamma(x,t)x^T]B_1^T - \mathbb{E}[\gamma(x,t)]B_3(t)^T - B_1\mathbb{E}[\gamma(x,t)x^T]^T \]
\[ - B_3(t)\mathbb{E}[\gamma(x,t)]^T - \Lambda(\mathbb{E}[x], \mathbb{E}[\gamma(x,t)], t, 0), \]
\[ A = B_2^T. \]

From Assumption 3.3, we have that the matrix \( B_2 \) is Hurwitz, and therefore, there exists a unique solution for \( \mathbb{E}[zz^T] \) in the equation (B.14). Thus, to prove that the solution to (B.14) is in the form of

\[ \mathbb{E}[zz^T] = \mathbb{E}[\gamma(x,t)\gamma(x,t)^T + g(x,t)g(x,t)^T] \tag{A.11} \]

given by (3.14) for the case where \( P_z = 2 \) we substitute (A.11) in (B.14), which yields

\[
\begin{align*}
\mathbb{E}[\gamma(x,t)\gamma(x,t)^T]B_2^T + \mathbb{E}[g(x,t)g(x,t)^T]B_2^T \\
+ B_2\mathbb{E}[\gamma(x,t)\gamma(x,t)^T] + B_2\mathbb{E}[g(x,t)g(x,t)^T] = \\
- \mathbb{E}[\gamma(x,t)x^T]B_1^T - \mathbb{E}[\gamma(x,t)]B_3(t)^T - B_1\mathbb{E}[\gamma(x,t)x^T]^T \\
- B_3(t)\mathbb{E}[\gamma(x,t)]^T - \Lambda(\mathbb{E}[x], \mathbb{E}[\gamma(x,t)], t, 0).
\end{align*}
\]

Simplifying further using the linearity of the expectation operator and the function \( \gamma(x,t) \), and noting that \( B_1x + B_3(t) = -B_2\gamma(x,t) \) from the expression for \( \gamma(x,t) \) in equation (3.4), we have that

\[
\begin{align*}
\mathbb{E}[\gamma(x,t)\gamma(x,t)^T]B_2^T + \mathbb{E}[g(x,t)g(x,t)^T]B_2^T \\
+ B_2\mathbb{E}[\gamma(x,t)\gamma(x,t)^T] + B_2\mathbb{E}[g(x,t)g(x,t)^T] = \\
\mathbb{E}[\gamma(x,t)\gamma(x,t)^T]B_2^T + B_2\mathbb{E}[\gamma(x,t)\gamma(x,t)^T] \\
- \Lambda(\mathbb{E}[x], \gamma(\mathbb{E}[x], t), t, 0).
\end{align*}
\]

Canceling the common terms on both sides yields the expression
\[ \mathbb{E}[g(x,t)g(x,t)^T]B_2^2 + B_2\mathbb{E}[g(x,t)g(x,t)^T] = -\Lambda(\mathbb{E}[x], \gamma(\mathbb{E}[x], t), t, 0), \]

which is satisfied by the definition of the function \( g(x,t) \) in (3.7). Thus, we have that setting \( \epsilon = 0 \) in the moments of the full system yields the equation (3.14) for \( P_z = 2 \), i.e. for the second order moments of \( z \).

### A.4 Proof of Theorem 3.1

From the commutative diagram in Lemma 3.1, it follows that setting \( \epsilon = 0 \) in the moment dynamics of the full system yields the moment dynamics of the reduced-order system for up to second order moments of the fast variable and for all moments of the slow variable. Therefore, we can apply Tikhonov’s theorem to the moment dynamics of the full system in (3.8)–(3.10) to obtain the result (3.15)–(3.16).

We first prove that the assumptions of the Tikhonov’s theorem are satisfied. To this end, we begin by considering the assumption on the global exponential stability of the boundary layer dynamics of the system (3.8)–(3.10). From equation (3.8), it follows that the fast variables that appear in (3.8) are of the form \( \mathbb{E}[z_ix^{(j)}] \) for \( j \in G_{ns}^{P-1} \). Thus, we define the vector \( b_j = [b_{ij}, \ldots, b_{mj}]^T \) for the boundary layer variable for \( \mathbb{E}[z_ix^{(j)}] \) where \( b_{ij} = \mathbb{E}[z_ix^{(j)}] - \mathbb{E}[\gamma_i(x,t)x^{(j)}] \) for \( j \in G_{ns}^{P-1} \) and \( i = 1, \ldots, m \). Furthermore, in Theorem 3.1, we only consider up to the second order moments of the fast variable. Thus, we define the matrix \( V \) as the boundary layer variable for \( \mathbb{E}[z_lz_k] \) where the elements of \( V \) are given by \( v_{lk} = \mathbb{E}[z_lz_k] - \mathbb{E}\left[\gamma_l(x,t)\gamma_k(x,t) + \sum_{h=1}^{d_s+d_f} g(x,t)_{lh}g(x,t)_{kh}\right] \) for \( l, k = 1, \ldots, m \). Then the dynamics of the variable \( b_{ij} \) and \( v_{lk} \) are given by

\[
\frac{db_{ij}}{dt} = \frac{d\mathbb{E}[z_ix^{(j)}]}{dt} - \frac{d\mathbb{E}[\gamma_i(x,t)x^{(j)}]}{dt},
\]

\[
\frac{dv_{lk}}{dt} = \frac{d\mathbb{E}[z_lz_k]}{dt} - \frac{d\mathbb{E}\left[\gamma_l(x,t)\gamma_k(x,t) + \sum_{h=1}^{d_s+d_f} g(x,t)_{lh}g(x,t)_{kh}\right]}{dt}.
\]
Let $\tau = t/\epsilon$ be the time variable in the fast time-scale. Then we have that

$$
\frac{dB_{ij}}{d\tau} = \epsilon \frac{\partial E[z_i x^{(j)}]}{\partial t} - \epsilon \frac{\partial E[\gamma(x, t) x^{(j)}]}{\partial t},
$$

$$
\frac{dv_{lk}}{d\tau} = \epsilon \frac{\partial E[z_l z_k]}{\partial t} - \epsilon \frac{\partial E[\gamma(x, t) \gamma_k(x, t) + \sum_{h=1}^{d_x + d_f} g(x, t) u_h g(x, t) k_h]}{\partial t}.
$$

Since from (3.4) we have that $\gamma_i(x, t)$ is a linear function of $x$, and since $j \in \mathcal{G}^{P-1}_{n_s}$ we have that $x^{(j)}$ contains moments of order up to $P - 1$. Therefore, it follows that $\gamma_i(x, t) x^{(j)}$ can be written in terms of $P^{th}$ or lower order moments of $x$ and $\gamma_i(x, t) \gamma_k(x, t)$ consists of up to second order moments of $x$. Furthermore, from (3.7) we have that $g(x, t) g(x, t)^T$ is a matrix whose elements are linear functions of $x$. Therefore, for appropriate functions $Q_k(t)$ for $k \in \mathcal{G}^P_{n_s}, Z_r(t)$ for $r \in \mathcal{G}^2_{n_s}$ and employing the linearity of the differential operator, we can write

$$
\frac{dB_{ij}}{d\tau} = \frac{\partial E[z_i x^{(j)}]}{\partial t} - \epsilon \sum_{k \in \mathcal{G}^P_{n_s}} Q_k(t) \frac{\partial E[x^{(k)}]}{\partial t},
$$

$$
\frac{dv_{lk}}{d\tau} = \frac{\partial E[z_l z_k]}{\partial t} - \epsilon \sum_{r \in \mathcal{G}^2} Z_r(t) \frac{\partial E[x^{(r)}]}{\partial t}.
$$

Substituting from (3.8) and using the expansions of $\partial E[z_i x^{(j)}]/\partial t, \partial E[z_l z_k]/\partial t$, (see proof of Claim 3.2), yields

$$
\frac{dB_{ij}}{d\tau} = \mathbb{E} \left[ f_{x_i}(x, z, t, \epsilon) x^{(j)} \right] + \epsilon \sum_{i=1}^{n} j_i \mathbb{E} \left[ f_{x_i}(x, z, t) z_i x_1^{j_1} \ldots x_i^{j_{n_s}} \ldots x_n^{j_{n_s}} \right]
$$

$$
+ \epsilon \frac{1}{2} \sum_{p=1}^{n} j_p (j_p - 1) \mathbb{E} \left[ \phi_{pp}(x, z, t) z_i x_1^{j_1} \ldots x_p^{j_{p-2}} \ldots x_n^{j_{n_s}} \right]
$$

$$
+ \sqrt{\epsilon} \sum_{m=1}^{n} j_{n_f} \mathbb{E} \left[ \theta_{im}(x, z, t, \epsilon) x_1^{j_1} \ldots x_{n_f}^{j_{n_f}} \ldots x_n^{j_{n_s}} \right]
$$

$$
+ \epsilon \sum_{m=2}^{n} \sum_{p=1}^{m-1} j_{n_f} j_p \mathbb{E} \left[ \phi_{pm}(x, z, t) z_i x_1^{j_1} \ldots x_{j_{p-1}}^{j_{p-1}} \ldots x_p^{j_{n_f}} \ldots x_n^{j_{n_s}} \right]
$$

$$
- \epsilon \sum_{s=0}^{P} \sum_{k \in \mathcal{K}_s} Q_k(t) \left( \sum_{i \in \mathcal{N}_P} C_{1i}(t) \mathbb{E}[x^{(i)}] + \sum_{l \in \mathcal{C}_1} \sum_{j \in \mathcal{N}_{P-1}} C_{2lj}(t) \mathbb{E}[(\gamma(x, t)^{(l)} x^{(j)})] \right). \quad (A.12)
$$
\[
\frac{dv_{lk}}{dt} = E[z_1 f_{zk}(x, z, t, \epsilon)] + E[z_k f_{zl}(x, z, t, \epsilon)] + E[\lambda_{ik}(x, z, t, \epsilon)] - \epsilon \sum_{r \in \mathbb{N}^{2}} Z_r(t) \frac{dE[x^{(r)}]}{dt}.
\]

(A.13)

where \( \phi_{ij}(x, z, t) \), \( \theta_{ij}(x, z, t, \epsilon) \), \( \lambda_{ik}(x, z, t, \epsilon) \) for \( i, j = 1, \ldots n \) and \( l, k = 1, \ldots, m \) are the elements of the matrices \( \Phi(x, z, t) \), \( \Theta(x, z, t, \epsilon) \) and \( \Lambda(x, z, t, \epsilon) \) given in Assumption 3.2.

We obtain the boundary layer system for the vector \( b_j \), by setting \( \epsilon = 0 \) in the equation (A.12), with \( E[zx^{(j)}] = b_j + E[\gamma(x, t)x^{(j)}] \). Under Assumptions 3.1 - 3.2, the functions \( f_x(x, z, t) \), \( f_z(x, z, t, \epsilon) \), \( \phi_{ij}(x, z, t) \), \( \theta_{ij}(x, z, t, \epsilon) \), \( Q_k(t) \), \( C_{1l}(t) \), \( C_{2ji}(t) \) are continuous and therefore are bounded on any compact interval \( t = [0, t_1] \). We also have that \( \theta_{ij}(x, z, t, 0) < \infty \), from Assumption 3.2. Additionally, as the system (3.8)-(3.10) is linear, the solutions \( E[x^{(i)}] \) and \( E[\gamma(x, t)x^{(j)}] \) exist and are bounded on any compact interval \( t = [0, t_1] \). Therefore, the boundary layer dynamics \( b_j \) are given by

\[
\frac{db_j}{dt} = E[f_x(x, z, t, 0)x^{(j)}] \bigg|_{E[zx^{(j)}]=b_j+E[\gamma(x, t)x^{(j)}]}
\]

Together with Assumption 3.1 and the definition of \( \gamma(x, t) \) in equation (3.4), we then obtain

\[
\frac{db_j}{dr} = B_1 E[xx^{(j)}] + B_2 (b_j + E[\gamma(x, t)x^{(j)}]) + B_3(t)E[x^{(j)}]
\]

(A.14)

In order to obtain the boundary layer dynamics for \( V \), we first represent the elements of \( b_j \) where \( j = (k_1, \ldots, k_n) \) \( \in \mathcal{G}^{P-1}_{n_x} \) for the case \( P = 1 \) as a vector \( d = E[z] - E[\gamma(x, t)] \) and for the case \( P = 2 \) as a matrix \( E = E[zx^T] - E[\gamma(x, t)x^T] \). Then, the boundary layer dynamics for the matrix \( V \) can be obtained by setting \( \epsilon = 0 \) in (A.13) and taking \( E[zx^T] = E + E[\gamma(x, t)x^T] \) and \( E[z] = d + E[\gamma(x, t)] \), which yields

\[
\frac{dV}{dr} = VB_2^T + B_2 V^T + EB_1^T + B_1 E^T + B_3(t)d^T + dB_3(t)^T.
\]

(A.15)

Under Assumption 3.3, we have that the origin is a globally exponential stable equilibrium point of the boundary layer dynamics \( b_j \) in (A.14). Next, to determine the stability of the boundary layer dynamics \( V \), we consider the solution of (A.15)
for $V$ given by [51]

$$V(\tau) = e^{B_2 \tau} V(0) e^{B_2 T} + \int_0^T e^{B_2 (\tau - \tau)} (E(\tau) B_1^T + B_1 E(\tau) T + B_3(t) d(\tau) + d(\tau) B_1(t)^T e^{B_2 (\tau - \tau) T}) d\tau.$$

Then, considering the solutions for $E$ and $d$, which can be obtained from (A.14), and are in the form $E(\tau) = E(0) e^{B_2 \tau}$ and $d(\tau) = d(0) e^{B_2 \tau}$, and using that $B_2$ is Hurwitz under Assumption 3.3, it follows that there exists positive constants $C_1$ and $r_1$ such that $\|V(\tau)\|_F \leq C_1(\|d(0)\|_F + \|E(0)\|_F + \|V(0)\|_F) e^{-r_1 \tau}$, where $\|.\|_F$ denotes the Frobenius norm. Then, taking $Y = [d \mid E \mid V]$, and considering the exponential stability of $E$ and $d$, we can write $\|Y(\tau)\|_F \leq C\|Y(0)\|_F e^{-r \tau}$ for positive constants $C$ and $r$. Thus, we have that the origin is a globally exponentially stable equilibrium point of the boundary layer dynamics $V$.

Furthermore, we ensure that the additional assumptions of the Tikhonov's theorem also hold. We have that $C_{1i}(t), C_{2i}(t)$ are continuous functions with respect to time and that the functions $D_{1a}(t, \epsilon), D_{2bc}(t, \epsilon), F_{1a}(t, \epsilon), F_{2a}(t, \epsilon), F_{3qrs}(t, \epsilon)$ in (3.8)-(3.10) and their partial derivatives with respect to $t$ and $\epsilon$ are continuous, from Claim 3.2. Due to the linearity of the function $\gamma(x, t)$ and $g(x, t)g(x, t)^T$, we have that the function $E[\gamma(x, t)x^{(j)}(\cdot)]$ for $j \in G_{n_2}^{P-1}$ has continuous first partial derivatives with respect to its arguments $E[x^{(k)}(\cdot)]$ for $k \in G_{d, o}^P$ and the function $E[\gamma(x, t)\gamma(x, t)^T + g(x, t)g(x, t)^T]$ has continuous first partial derivatives with respect to its arguments $E[x^{(k)}(\cdot)]$ for $k \in G_{d, o}^2$. Furthermore, we have that the system (3.11) has a unique solution on a compact time interval $t \in [0, t_1]$, due to its linearity. Hence, the assumptions of the Tikhonov's theorem on a finite time interval are satisfied and applying the theorem to the moment dynamics of the full system in (3.8)-(3.10), yields the desired result in (3.15)-(3.16).
Appendix B

Proofs pertaining to Chapter 4

B.1 Proof of Claim 4.1

Applying the coordinate transformation \( x = T_x v, z = T_z v \) to equation (4.1), with \( \tilde{a}(v, t) = [\tilde{a}_s(v, t), (1/\epsilon)\tilde{a}_f(v, t)]^T \) and \( q = [q_1, \ldots, q_{m_s}, q_{m_s+1}, \ldots, q_{m_s+m_f}] \), with \( v = T^{-1}[x^T, z^T]^T \) we have

\[
\dot{x} = T_x f(T^{-1}[x^T, z^T]^T, t) = T_x \sum_{i=1}^{m_s} q_i \dot{a}_{si}(T^{-1}[x^T, z^T]^T, t) \\
+ T_x \sum_{i=m_s+1}^{m_s+m_f} q_i (1/\epsilon) \dot{a}_{fi}(T^{-1}[x^T, z^T]^T, t) \\
= f_x(x, z, t), \tag{B.1}
\]

\[
\dot{z} = T_z f(T^{-1}[x^T, z^T]^T, t) = T_z \sum_{i=1}^{m_s} q_i \dot{a}_{si}(T^{-1}[x^T, z^T]^T, t) \\
+ T_z \sum_{i=m_s+1}^{m_s+m_f} q_i (1/\epsilon) \dot{a}_{fi}(T^{-1}[x^T, z^T]^T, t) \\
= \frac{1}{\epsilon} f_z(x, z, t, \epsilon). \tag{B.2}
\]

Thus, from equation (B.1), it follows that \( T_x q_i = 0 \) for \( i = m_s + 1, \ldots, m_s + m_f \). Thus, the fast reactions do not appear in the slow dynamics, however, the slow reactions can appear in the fast dynamics.
Applying the coordinate transformation $\psi_x = T_x \xi$, $\psi_z = T_z \xi$, to equation (4.2), we have that

\[
\dot{\psi}_x = T_x [A(v, t) \xi] + T_x \sigma(v, t) \Gamma,
\]
\[
\dot{\psi}_z = T_z [A(v, t) \xi] + T_z \sigma(v, t) \Gamma.
\]

Since $A(v, t) = \frac{\partial f(x, t)}{\partial v}$ and $v = T^{-1}[x^T, z^T]^T$, using the chain rule we can write

\[
\dot{\psi}_x = T_x \left[ \frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial z} \frac{\partial z}{\partial v} \right] \xi
\]
\[
+ T_x \left[ q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma,
\]
\[
\dot{\psi}_z = T_z \left[ \frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f(T^{-1}[x^T, z^T]^T, t)}{\partial z} \frac{\partial z}{\partial v} \right] \xi
\]
\[
+ T_z \left[ q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma.
\]

Using the linearity of the differentiation operator and the transformation $x = T_x v$, $z = T_z v$, we obtain

\[
\dot{\psi}_x = \left[ \frac{\partial T_x f(T^{-1}[x^T, z^T]^T, t)}{\partial x} T_x + \frac{\partial T_z f(T^{-1}[x^T, z^T]^T, t)}{\partial z} T_z \right] \xi
\]
\[
+ T_x \left[ q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma,
\]
\[
\dot{\psi}_z = \left[ \frac{\partial T_z f(T^{-1}[x^T, z^T]^T, t)}{\partial x} T_x + \frac{\partial T_z f(T^{-1}[x^T, z^T]^T, t)}{\partial z} T_z \right] \xi
\]
\[
+ T_z \left[ q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma.
\]

From (B.1)–(B.2), we have that $T_x f(T^{-1}[x^T, z^T]^T, t) = f_x(x, z, t)$ and $T_z f(T^{-1}[x^T, z^T]^T, t) = \frac{1}{\epsilon} \tilde{a}_s(T^{-1}[x^T, z^T]^T, t)$, $(1/\epsilon) \tilde{a}_f(T^{-1}[x^T, z^T]^T, t)$, $T_x \tilde{a}_s(T^{-1}[x^T, z^T]^T, t)$, $T_z \tilde{a}_f(T^{-1}[x^T, z^T]^T, t)$, $T_x \sqrt{\tilde{a}_s(T^{-1}[x^T, z^T]^T, t)}$, $T_z \sqrt{\tilde{a}_f(T^{-1}[x^T, z^T]^T, t)}$, we have

\[
\dot{\psi}_x = \frac{\partial f_x(x, z, t)}{\partial x} \psi_x + \frac{\partial f_x(x, z, t)}{\partial z} \psi_z + T_x \left[ q_1 \sqrt{\tilde{a}_1(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_m \sqrt{\tilde{a}_m(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_x
\]
\[
+ T_x \left[ q_{m+1} \sqrt{\frac{1}{\epsilon} \tilde{a}_f(T^{-1}[x^T, z^T]^T, t)}, \ldots, q_{m+m_f} \sqrt{\frac{1}{\epsilon} \tilde{a}_f(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_f, \quad (B.3)
\]

144
\[ \psi_z = \frac{\partial^2 f_z(x, z, t, \epsilon)}{\partial x} \psi_x + \frac{\partial^2 f_z(x, z, t, \epsilon)}{\partial z} \psi_z + \]
\[ T_z \left[ q_1 \sqrt{\hat{a}_z(T^{-1}[x^T, z^T]^T, t), \ldots, q_{m_z} \sqrt{\hat{a}_{m_z}(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_x \]
\[ + T_z \left[ q_{m_z+1} \sqrt{\frac{1}{\epsilon} \hat{a}_{f_1}(T^{-1}[x^T, z^T]^T, t), \ldots, q_{m_z+m_f} \sqrt{\frac{1}{\epsilon} \hat{a}_{f_{m_f}}(T^{-1}[x^T, z^T]^T, t)} \right] \Gamma_f, \quad (B.4) \]

where \( \Gamma = [\Gamma_x^T, \Gamma_f^T]^T \). From (B.1) we have that, \( T_z q_i = 0 \) for \( i = m_z + 1, \ldots, m_z + m_f \).

Then, multiplying (B.4) by \( \epsilon \), and taking \( \Gamma_z = [\Gamma_z^T, \Gamma_f^T]^T \), we can write the system (B.3)–(B.4) in the form of system (4.5)–(4.6).

### B.2 Proof of Claim 4.2

We express the equations (4.9) - (4.10) in the form

\[ \dot{\psi}_x = S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z + [\sigma_x(x, z, t) 0] \Gamma_z, \]
\[ \epsilon \dot{\psi}_z = F_x(x, z, t, \epsilon)\psi_x + F_z(x, z, t, \epsilon)\psi_z + \sigma_z(x, z, t, \epsilon) \Gamma_z, \]

where \([ \sigma_x(x, z, t) 0 ] \in \mathbb{R}^{n \times (m_z+m_f)}\). Then, as the \( x \) and \( z \) are deterministic we use the linearity of the expectation operator to derive the dynamics for the first moments as

\[ \frac{d\mathbb{E}[\psi_x]}{dt} = S_x(x, z, t)\mathbb{E}[\psi_x] + S_z(x, z, t)\mathbb{E}[\psi_z], \quad (B.5) \]
\[ \frac{d\mathbb{E}[\psi_z]}{dt} = \frac{1}{\epsilon} F_x(x, z, t, \epsilon)\mathbb{E}[\psi_x] + \frac{1}{\epsilon} F_z(x, z, t, \epsilon)\mathbb{E}[\psi_x]. \quad (B.6) \]

Similarly, Proposition III.1 in Bence et al.\[59\] can be used to write the second moment dynamics as

\[ \frac{d}{dt} \mathbb{E} \begin{bmatrix} \psi_x \psi_x^T & \psi_x \psi_z^T \\ \psi_z \psi_x^T & \psi_z \psi_z^T \end{bmatrix} = \]
\[ \begin{bmatrix} \psi_x (S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z)^T & \frac{1}{\epsilon} \psi_x (F_x(x, z, t, \epsilon)\psi_x + F_z(x, z, t, \epsilon)\psi_z)^T \\ \psi_z (S_x(x, z, t)\psi_x + S_z(x, z, t)\psi_z)^T & \frac{1}{\epsilon} \psi_z (F_x(x, z, t, \epsilon)\psi_x + F_z(x, z, t, \epsilon)\psi_z)^T \end{bmatrix} \]
Then, summing the corresponding entries of the matrices in equation (B.7) and using the linearity of the expectation operator, the equations (B.5)-(B.7) can be written in the form (4.18)-(4.22). We have that \( \mathbb{E}[\psi_x^T] = (\mathbb{E}[\psi_x^T])^T \), and thus, we do not consider the dynamics of the variable \( \mathbb{E}[\psi_x^T] \) in the equations (4.18)-(4.22). Furthermore, the initial conditions \( \psi_{x0} \) and \( \psi_{z0} \) are deterministic, which yields \( \mathbb{E}[\psi_x(0)] = \psi_{x0}, \mathbb{E}[\psi_x^T(0)] = \psi_{x0}^T, \mathbb{E}[\psi_x^T(0)] = \psi_{z0}, \mathbb{E}[\psi_x^T(0)] = \psi_{z0}^T \).

### B.3 Proof of Claim 4.3

Using the Proposition III.1 in Bence et al. [59], the first and second moment dynamics of \( \psi_x \) in (4.14) can be written in the form

\[
\begin{align*}
\frac{d\mathbb{E}[\psi_x]}{dt} &= \mathbb{E}[S(\bar{x}, t)\psi_x], \\
\frac{d\mathbb{E}[\psi_x\psi_x^T]}{dt} &= \mathbb{E}[S(\bar{x}, t)\bar{\psi}_x\bar{\psi}_x^T] + \mathbb{E}[\bar{\psi}_x(\psi_x^T S(\bar{x}, t))]
+ \sigma_x(\bar{x}, \gamma_1(\bar{x}, t), t)\sigma_x(\bar{x}, \gamma_1(\bar{x}, t), t)^T.
\end{align*}
\]

We have that the dynamics of \( \bar{x} \) in (4.13) are deterministic. Therefore, using the linearity of the expectation operator, the moment dynamics of \( \psi_x \) in (4.14) can be written in the form of (4.23)-(4.24).

In order to derive the moments of the variable \( \psi_x \), we take the expected value of equation (4.15), which yields \( \mathbb{E}[\psi_x] = \mathbb{E}[\gamma_2(\bar{x}, t)\psi_x] \) as the elements of the vector \( N(0, 1) \) are normal random variable with zero mean. Since \( \bar{x} \) is deterministic, we have
that $E[\tilde{\psi}_z] = \gamma_2(\tilde{x}, t)E[\tilde{\psi}_x]$. Calculating the second moment of $\tilde{\psi}_x$, we obtain

$$E[\tilde{\psi}_x\tilde{\psi}_x^T] = E[(\gamma_2(\tilde{x}, t)\psi_x + g(\tilde{x}, t)N(0, 1))(\gamma_2(\tilde{x}, t)\psi_x + g(\tilde{x}, t)N(0, 1))^T].$$

Expanding further and using the fact that $N(0, 1)$ is independent of $\tilde{\psi}_x$, we have

$$E[\tilde{\psi}_x\tilde{\psi}_x^T] = \gamma_2(\tilde{x}, t)E[\psi_x\psi_x^T]\gamma_2(\tilde{x}, t)^T + g(\tilde{x}, t)E[N(0, 1)]^Tg(\tilde{x}, t)^T + g(\tilde{x}, t)E[N(0, 1)N(0, 1)^T]g(\tilde{x}, t)^T.$$

Since $N(0, 1)$ is a vector of standard normal random variables, we have that $E[N(0, 1)] = 0$ and $E[N(0, 1)N(0, 1)^T] = I_{d\times d}$ where $I_{d\times d}$ is an $d \times d$ identity matrix. Thus, we have that $E[\tilde{\psi}_x\tilde{\psi}_x^T] = \gamma_2(\tilde{x}, t)E[\psi_x\psi_x^T]\gamma_2(\tilde{x}, t)^T + g(\tilde{x}, t)g(\tilde{x}, t)^T$.

**B.4 Proof of Claim 4.4**

Setting $\epsilon = 0$, we have

$$0 = f_z(x, z, t, 0), \quad (B.8)$$

$$0 = F_x(x, z, t, 0)E[\psi_x] + F_z(x, z, t, 0)E[\psi_z], \quad (B.9)$$

$$0 = F_x(x, z, t, 0)E[\psi_x\psi_x^T] + F_z(x, z, t, 0)E[\psi_x\psi_x^T], \quad (B.10)$$

$$0 = F_x(x, z, t, 0)E[\psi_x\psi_x^T] + F_z(x, z, t, 0)E[\psi_x\psi_x^T] + E[\psi_z\psi_z^T]F_x(x, z, t, 0)^T + E[\psi_z\psi_z^T]F_z(x, z, t, 0)^T + \Lambda(x, z, t, 0). \quad (B.11)$$

Under Assumption 4.3, there exists an isolated real root $z = \gamma_1(x, t)$ for equation (B.8). Thus, the corresponding unique solutions to equations (B.9)-(B.10) are given
by

\[ \mathbb{E}[\psi_z] = -F_z(x, \gamma_1(x, t), t, 0)^{-1}(F_z(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_z]) \]
\[ = \gamma_2(x, t)\mathbb{E}[\psi_z], \quad (B.12) \]
\[ \mathbb{E}[\psi_z \psi_z^T] = -F_z(x, \gamma_1(x, t), t, 0)^{-1}(F_z(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_z \psi_z^T]) \]
\[ = \gamma_2(x, t)\mathbb{E}[\psi_z \psi_z^T]. \quad (B.13) \]

Substituting \( z = \gamma_1(x, t) \) and equations (B.12)–(B.13) in (4.7) and the moment equations (4.18)–(4.22), yields the dynamics of \( \bar{x} \) given by (4.13) and the moment dynamics of the variable \( \bar{\psi}_z \) reduced system given by (4.23)–(4.24).

Furthermore, we have that the equation (B.12) is equivalent to the first moment of the variable \( \bar{\psi}_z \) of the reduced system given by (4.25). Next, in order to solve equation (B.11), we substitute (B.12)–(B.13) in (B.11), which yields

\[ F_z(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_z \psi_z^T] + \mathbb{E}[\psi_z \psi_z^T]F_z(x, \gamma_1(x, t), t, 0)^T \]
\[ = -F_z(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_z \psi_z^T]\gamma_2(x, t)^T \]
\[ - \gamma_2(x, t)\mathbb{E}[\psi_z \psi_z^T]F_z(x, \gamma_1(x, t), t, 0)^T \]
\[ - \Lambda(x, \gamma_1(x, t), t, 0), \quad (B.14) \]

which is in the form of a Lyapunov equation \( AP + PA^T = -Q \) with

\[ A = F_z(x, \gamma_1(x, t), t, 0), \]
\[ P = \mathbb{E}[\psi_z \psi_z^T], \]
\[ Q = F_z(x, \gamma_1(x, t), t, 0)\mathbb{E}[\psi_z \psi_z^T]\gamma_2(x, t)^T \]
\[ + \gamma_2(x, t)\mathbb{E}[\psi_z \psi_z^T]F_z(x, \gamma_1(x, t), t, 0)^T + \Lambda(x, \gamma_1(x, t), t, 0) \]

Under Assumption 4.3, the matrix \( F_z(x, \gamma_1(x, t), t, 0) \) is Hurwitz for all \( x \) and \( t \) and thus, the equation (B.14) has a unique solution \( \mathbb{E}[\psi_z \psi_z^T] = h(x, \mathbb{E}[\psi_z \psi_z^T], t) \). In order to prove that \( h(x, \mathbb{E}[\psi_z \psi_z^T], t) = \gamma_2(x, t)\mathbb{E}[\psi_z \psi_z^T]\gamma_2(x, t)^T + \Lambda(x, \gamma_1(x, t), t, 0)^T \) given by the second moment of the \( \bar{\psi}_z \) of the reduced system in (4.26), we substitute \( \mathbb{E}[\psi_z \psi_z^T] = \)
\[ \gamma_2(x, t)E[\psi_x \psi_x^T] \gamma_2(x, t)^T + g(x, t)g(x, t)^T, \] which yields

\[
F_z(x, \gamma_1(x, t), t, 0) \gamma_2(x, t)E[\psi_x \psi_x^T] \gamma_2(x, t)^T + F_z(x, \gamma_1(x, t), t, 0)g(x, t)g(x, t)^T + \gamma_2(x, t)E[\psi_x \psi_x^T] \gamma_2(x, t)^T F_z(x, \gamma_1(x, t), t, 0)^T + g(x, t)g(x, t)^T F_z(x, \gamma_1(x, t), t, 0)^T
\]

\[
= -F_z(x, \gamma_1(x, t), t, 0)E[\psi_x \psi_x^T] \gamma_2(x, t)^T - \gamma_2(x, t)E[\psi_x \psi_x^T] F_z(x, \gamma_1(x, t), t, e)^T
\]

\[- \Lambda(x, \gamma_1(x, t), t, 0)\]

Then, canceling the common terms on both sides we obtain

\[
F_z(x, \gamma_1(x, t), t, 0)g(x, t)g(x, t)^T + g(x, t)g(x, t)^T F_z(x, \gamma_1(x, t), t, 0)^T = -\Lambda(x, \gamma_1(x, t), t, 0)
\]

which satisfies the equation (4.17) in the definition of the reduced fast system.

**B.5 Proof of Theorem 4.1**

It can be seen from the commutative diagram in Lemma 4.1 that setting \( \epsilon = 0 \) in the moment equations of the full system together with the dynamics of \( x \) and \( z \) yields the moment equations of the reduced system. Thus, as the moment dynamics are deterministic, we can apply the Tikhonov’s theorem to the moments of the full system together with the dynamics of \( x \) and \( z \) to prove the results given in Theorem 4.1. Towards this end, we first prove that the assumptions of the Tikhonov’s theorem are satisfied. We first consider the boundary layer dynamics of the moment dynamics (4.18)-(4.22), where we define the boundary layer variables as

\[
b_1 = z - \gamma_1(x, t), \quad (B.15)
\]

\[
b_2 = E[\psi_x] - \gamma_2(x, t)E[\psi_x], \quad (B.16)
\]

\[
b_3 = E[\psi_x \psi_x^T] - \gamma_2(x, t)E[\psi_x \psi_x^T], \quad (B.17)
\]

\[
b_4 = E[\psi_x \psi_x^T] - (\gamma_2(x, t)E[\psi_x \psi_x^T] \gamma_2(x, t)^T + g(x, t)g(x, t)^T). \quad (B.18)
\]
Define $\gamma_3(x, \mathbb{E}[\psi_x^T], t) = \gamma_2(x, t)\mathbb{E}[\psi_x^T]\gamma_2(x, t)^T + g(x, t)g(x, t)^T$. Then, we have that the dynamics of the boundary layer variables are given by

$$\begin{align*}
\frac{db_1}{dt} &= \frac{dz}{dt} - \frac{d\gamma_1(x, t)}{dt}, \\
\frac{db_2}{dt} &= \frac{d\mathbb{E}[\psi_x]}{dt} - \frac{d\gamma_2(x, t)\mathbb{E}[\psi_x]}{dt}, \\
\frac{db_3}{dt} &= \frac{d\mathbb{E}[\psi_x^T]}{dt} - \frac{d\gamma_2(x, t)\mathbb{E}[\psi_x^T]}{dt}, \\
\frac{db_4}{dt} &= \frac{d\mathbb{E}[\psi_x^T]}{dt} - \frac{d\gamma_3(x, \mathbb{E}[\psi_x^T], t)}{dt}.
\end{align*}$$

Taking $\tau = t/\epsilon$ to be the time variable in the fast time-scale we have that

$$\begin{align*}
\frac{db_1}{d\tau} &= \epsilon \frac{dz}{dt} - \epsilon \frac{\partial\gamma_1(x, t)}{\partial t} - \epsilon \frac{\partial\gamma_1(x, t)}{\partial x} \frac{dx}{dt}, \\
\frac{db_2}{d\tau} &= \epsilon \frac{d\mathbb{E}[\psi_x]}{dt} - \epsilon \frac{\mathbb{E}[\psi_x^T] \partial\gamma_2(x, t)}{\partial t} - \epsilon \frac{\mathbb{E}[\psi_x]}{\partial x} \frac{d\mathbb{E}[\psi_x]}{dt}, \\
\frac{db_3}{d\tau} &= \epsilon \frac{d\mathbb{E}[\psi_x^T]}{dt} - \epsilon \frac{\mathbb{E}[\psi_x^T] \partial\gamma_2(x, t)}{\partial t} - \epsilon \frac{\mathbb{E}[\psi_x]}{\partial x} \frac{d\mathbb{E}[\psi_x^T]}{dt}, \\
\frac{db_4}{d\tau} &= \epsilon \frac{d\mathbb{E}[\psi_x^T]}{dt} - \epsilon \frac{\partial\gamma_3(x, \mathbb{E}[\psi_x^T], t)}{\partial t} \\
&\quad - \epsilon \frac{\partial\gamma_3(x, \mathbb{E}[\psi_x^T], t)}{\partial x} \frac{dx}{dt} \\
&\quad - \epsilon \frac{\partial\gamma_3(x, \mathbb{E}[\psi_x^T], t)}{\partial \mathbb{E}[\psi_x^T]} \frac{d\mathbb{E}[\psi_x^T]}{dt}.
\end{align*}$$

Then, substituting for the moment dynamics from equations (4.8), (4.20), (4.21) and (4.22) we obtain

$$\begin{align*}
\frac{db_1}{d\tau} &= f_1(x, z, t, \epsilon) - \epsilon \frac{\partial\gamma_1(x, t)}{\partial t} - \epsilon \frac{\partial\gamma_1(x, t)}{\partial x} \frac{dx}{dt}, \\
\frac{db_2}{d\tau} &= F_x(x, z, t, \epsilon)\mathbb{E}[\psi_x] + F_z(x, z, t, \epsilon)\mathbb{E}[\psi_z] \\
&\quad - \epsilon \mathbb{E}[\psi_x] \frac{\partial\gamma_2(x, t)}{\partial t} - \epsilon \mathbb{E}[\psi_x] \frac{\partial\gamma_2(x, t)}{\partial x} \frac{dx}{dt} \\
&\quad - \epsilon \frac{\partial\gamma_2(x, t)}{\partial \mathbb{E}[\psi_x]} \frac{d\mathbb{E}[\psi_x]}{dt},
\end{align*}$$

(B.19) (B.20)
where we take \( z = b_1 + \gamma_1(x, t) \) and \( \mathbb{E}[^\psi_x] = b_2 + \gamma_2(x, t)\mathbb{E}[\psi_x], \mathbb{E}[\psi_x^T] = b_3 + \gamma_2(x, t)\mathbb{E}[\psi_x^T] \) and \( \mathbb{E}[\psi_x^T] = b_4 + \gamma_3(x, \mathbb{E}[\psi_x^T], t) \). From Assumptions 4.1 - 4.3 we have that the functions \( f_x, f_z, \gamma_1, \gamma_2, \) and \( \gamma_3 \) are continuously differentiable and thus is bounded for a finite time interval \( t \in [0, t_1] \). Furthermore due to the linearity of the moment equations we have that \( \mathbb{E}[\psi_x] \) and \( \mathbb{E}[\psi_x^T] \) exist and are bounded for a finite time interval \( t \in [0, t_1] \). Thus, setting \( \epsilon = 0 \) in the equations (B.19)–(B.22), we obtain the boundary layer system given by

\[
\begin{align*}
\frac{db_1}{d\tau} &= f_x(x, b_1 + \gamma_1(x, t), t, 0), \quad \text{(B.23)} \\
\frac{db_2}{d\tau} &= F_x(x, b_1 + \gamma_1(x, t), t, 0)\mathbb{E}[\psi_x] \quad \text{(B.24)} \\
&\quad + F_x(x, b_1 + \gamma_1(x, t), t, 0)(b_2 + \gamma_2(x, t)\mathbb{E}[\psi_x]), \\
\frac{db_3}{d\tau} &= F_x(x, b_1 + \gamma_1(x, t), t, 0)\mathbb{E}[\psi_x^T] \\
&\quad + F_x(x, b_1 + \gamma_1(x, t), t, 0)(b_3 + \gamma_2(x, t)\mathbb{E}[\psi_x^T]), \quad \text{(B.25)} \\
\frac{db_4}{d\tau} &= F_x(x, b_1 + \gamma_1(x, t), t, 0)(b_3 + \gamma_2(x, t)\mathbb{E}[\psi_x^T])^T \\
&\quad + F_x(x, b_1 + \gamma_1(x, t), t, 0)(b_1 + \gamma_3(x, \mathbb{E}[\psi_x^T], t)) \\
&\quad + (b_3 + \gamma_2(x, t)\mathbb{E}[\psi_x^T])F_x(x, b_1 + \gamma_1(x, t), t, 0)^T
\end{align*}
\]
\[ + (b_4 + \gamma_3(x, E[\psi_x \psi_x^T], t))F_x(x, b_1 + \gamma_1(x, t), t, 0)^T \]
\[ + \Lambda(x, b_1 + \gamma_1(x, t), t), \]

(B.26)

Next, in order to prove that the origin is an exponentially stable equilibrium point of the boundary layer system, we linearize the system (B.23)-(B.26) about the origin. Towards this end, we first represent the matrix variable \( b_3 \) and \( b_4 \) in vector form. Let \( A \) be an \( m \times n \) matrix and let \( \text{vec}(A) = [a_{11}, \ldots, a_{m1}, \ldots, a_{1n}, \ldots, a_{mn}]^T \), where \( a_{ij} \) are the elements of the matrix \( A \). Then, considering the dynamics of \( \text{vec}(b_3) \) and \( \text{vec}(b_4) \), we use the Kronecker Product denoted by \( \otimes \) to obtain

\[
\begin{align*}
\frac{db_1}{dt} &= f_x(x, b_1 + \gamma_1(x, t), t, 0), \\
\frac{db_2}{dt} &= F_x(x, b_1 + \gamma_1(x, t), t, 0)E[\psi_x] \\
&+ F_z(x, b_1 + \gamma_1(x, t), t, 0)(b_2 + \gamma_2(x, t)E[\psi_x]), \\
\frac{d\text{vec}(b_3)}{dt} &= (I \otimes F_x(x, b_1 + \gamma_1(x, t), t, 0))\text{vec}(b_3) \\
&+ \text{vec}(g_2(b_1, x, \psi_x, t)), \\
\frac{d\text{vec}(b_4)}{dt} &= (I \otimes F_z(x, b_1 + \gamma_1(x, t), t, 0) \otimes I)\text{vec}(b_4) \\
&+ (I \otimes F_x(x, b_1 + \gamma_1(x, t), t, 0) \otimes I)\text{vec}(b_3) \\
&+ \text{vec}(g_3(b_1, x, \psi_x, t)),
\end{align*}
\]

where

\[
\begin{align*}
g_2(b_1, x, \psi_x, t) &= F_x(x, b_1 + \gamma_1(x, t), t, 0)E[\psi_x \psi_x^T] \\
&+ F_z(x, b_1 + \gamma_1(x, t), t, 0)\gamma_2(x, t)E[\psi_x \psi_x^T], \\
g_3(b_1, x, \psi_x, t) &= F_x(x, b_1 + \gamma_1(x, t), t, 0)E[\psi_x \psi_x^T]\gamma_2(x, t)^T \\
&+ F_z(x, b_1 + \gamma_1(x, t), t, 0)\gamma_3(x, E[\psi_x \psi_x^T], t)
\end{align*}
\]
\[ + \gamma_2(x, t) E[\psi_x \psi_x^T] F_x(x, b_1 + \gamma_1(x, t), t, 0)^T \]
\[ + \gamma_3(x, E[\psi_x \psi_x^T], t) F_x(x, b_1 + \gamma_1(x, t), t, 0)^T \]
\[ + \Lambda(x, b_1 + \gamma_1(x, t), t). \]

Then, considering the state vector \( e = [b_1, b_2, \text{vec}(b_3), \text{vec}(b_4)] \) and linearizing about \( \dot{e} = 0 \), we obtain the dynamics of \( \dot{e} = e - \dot{e} \) in the form

\[
\frac{d\dot{e}}{dt} = \begin{bmatrix}
J_{11} & 0 & 0 & 0 \\
J_{21} & J_{22} & 0 & 0 \\
J_{31} & 0 & J_{33} & 0 \\
J_{41} & 0 & J_{43} & J_{44}
\end{bmatrix} \dot{e}, \quad (B.27)
\]

where the diagonal entries are given by 
\[
J_{11} = \left. \frac{\partial f_x(x, b_1 + \gamma_1(x, t), t, 0)}{\partial b_1} \right|_{b_1=0}, \quad J_{22} = F_x(x, b_1 + \gamma_1(x, t), t, 0) \big|_{b_1=0}, \quad J_{33} = (I \otimes F_x(x, b_1 + \gamma_1(x, t), t, 0)) \big|_{b_1=0} \quad \text{and} \quad J_{44} = (F_x(x, b_1 + \gamma_1(x, t), t, 0) \otimes F_x(x, b_1 + \gamma_1(x, t), t, 0)) \big|_{b_1=0}
\]
where \( \otimes \) denotes the Kronecker sum and the lower diagonal entries \( J_{21}, J_{31}, J_{41}, J_{43} \) are appropriate functions.

Since the eigenvalues of a block triangular matrix are given by the union of eigenvalues of the diagonal blocks, we consider the eigenvalues of the diagonal entries. We have that 
\[
\frac{\partial f_x(x, b_1 + \gamma_1(x, t), t, 0)}{\partial b_1} \big|_{b_1=0} = \left. \frac{\partial f_x(x, z, t, 0)}{\partial z} \frac{\partial z}{\partial b_1} \right|_{z=\gamma_1(x, t)} = \left. \frac{\partial f_x(x, z, t, 0)}{\partial z} \right|_{z=\gamma_1(x, t)},
\]
which is Hurwitz from Assumption 4.3. Furthermore, we have that 
\[
F_x(x, z, t, \epsilon) = \left. \frac{\partial f_x(x, z, t, \epsilon)}{\partial z} \right|_{z=\gamma_1(x, t)}
\]
from the definition of the full system (4.7)–(4.10), and thus, \( F_x(x, b_1 + \gamma_1(x, t), t, 0) \big|_{b_1=0} \text{ is } F_x(x, b_1 + \gamma_1(x, t), t, 0) \big|_{b_1=0} \) is Hurwitz under Assumption 4.3. Thus, the diagonal term \( J_{22} \) is Hurwitz. Considering the eigenvalues of \( J_{33} \), we have that any eigenvalue of a Kronecker product of two matrices are given by the product of the eigenvalues of the individual matrices [53]. Thus, the eigenvalues of \( J_{33} \) are given by the eigenvalues of \( F_x(x, b_1 + \gamma_1(x, t), t, 0) \big|_{b_1=0} \), which is Hurwitz under Assumption 4.3. Next, we consider the diagonal term \( J_{44} \). From Theorem 13.16 in Laub [53] we have that any eigenvalue of \( J_{44} \) is given by the sum of two eigenvalues of \( F_x(x, b_1 + \gamma_1(x, t), t, 0) \). Since the matrix \( F_x \) is Hurwitz under Assumption 4.3, we have that all eigenvalues of \( J_{44} \) have negative real parts. Therefore, we have that
the eigenvalues of the linearized system (B.27) have negative real parts and thus the origin is an exponentially stable equilibrium point uniformly for $x$ and $t$.

Next, we prove that the initial conditions of the system

\[
\begin{align*}
\frac{db_1}{d\tau} &= f_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0), \\
\frac{db_2}{d\tau} &= F_z'(x_0, b_1 + \gamma_1(x_0, 0), 0, 0)\psi_{x_0} \\
&\quad + F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0)(b_2 + \gamma_2(x_0, 0)\psi_{x_0}), \\
\frac{d\text{vec}(b_3)}{d\tau} &= (I \otimes F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0))\text{vec}(b_3) \\
&\quad + \text{vec}(g_2(b_1, x_0, \psi_{x_0}, 0)), \\
\frac{d\text{vec}(b_4)}{d\tau} &= (I \otimes F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0) \\
&\quad + F_z(x_0, b_1 + \gamma_1(x_0, 0), 0, 0) \otimes I)\text{vec}(b_4) \\
&\quad + \text{vec}(g_3(b_1, b_3, x_0, \psi_{x_0}, 0)),
\end{align*}
\]

are in the region of attraction of the equilibrium point at the origin. From Assumption 4.3 we have that the initial condition $z_0$ is in the region of attraction of the equilibrium point $\mathbf{z} = \gamma_1(x_0, 0)$ of system $\frac{d\mathbf{z}}{d\tau} = f_z(x_0, \mathbf{z}, 0, 0)$. Thus, it follows that $z_0 - \gamma_1(x_0, 0)$ is in the region of attraction of equilibrium point $b_1 = 0$ for system (B.28). Furthermore, we have that the trajectory $b_1(\tau)$ that corresponds to the initial condition $z_0 - \gamma_1(x_0, 0)$ is bounded and $\lim_{\tau \to \infty} b_1(\tau) = 0$. Using this fact and the linearity of the system (B.29)–(B.31), we next prove that any trajectory of the system, starting with any initial condition, converges to zero as $\tau \to \infty$. Towards this end, we first define the vector $\mathbf{r} = [\hat{b}_2, \text{vec}(\hat{b}_3), \text{vec}(\hat{b}_4)]$ and write the system (B.29)–(B.31) in the form

\[
\frac{d\mathbf{r}}{d\tau} = \begin{bmatrix}
H_{11} & 0 & 0 \\
0 & H_{22} & 0 \\
0 & H_{32} & H_{33}
\end{bmatrix} \mathbf{r}
\]
+ \begin{bmatrix}
C_{11}(\tau) & 0 & 0 \\
0 & C_{22}(\tau) & 0 \\
0 & 0 & C_{32}(\tau) C_{33}(\tau)
\end{bmatrix} r + \begin{bmatrix}
0 \\
D_2(\tau) \\
D_3(\tau)
\end{bmatrix},

where

\begin{align*}
H_{11} &= F_z(x_0, \gamma_1(x_0, 0), 0, 0), \\
H_{22} &= I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0), \\
H_{32} &= I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0) + F_x(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I, \\
H_{33} &= I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0) + F_z(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I, \\
C_{11}(\tau) &= F_z(x_0, b_1(\tau) - \gamma_1(x_0, 0), 0, 0) - F_z(x_0, \gamma_1(x_0, 0), 0, 0), \\
C_{22}(\tau) &= I \otimes F_z(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) - I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0), \\
C_{32}(\tau) &= I \otimes F_z(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) + F_x(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) \otimes I \\
&\quad - I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0) - F_z(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I, \\
C_{33}(\tau) &= I \otimes F_z(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) + F_z(x_0, b_1(\tau) + \gamma_1(x_0, 0), 0, 0) \otimes I \\
&\quad - I \otimes F_z(x_0, \gamma_1(x_0, 0), 0, 0) - F_z(x_0, \gamma_1(x_0, 0), 0, 0) \otimes I, \\
D_2(\tau) &= \text{vec}(g_2(b_1(\tau), x_0, \psi_{x_0}, 0)), \\
D_3(\tau) &= \text{vec}(g_3(b_1(\tau), x_0, \psi_{x_0}, 0)).
\end{align*}

Then, we apply Lemmas 9.4–9.6 from Khalil[45] to show that \( r \) tends to zero as \( \tau \to \infty \) for any initial condition \( r(0) \). From Assumption 4.3, we have that the matrix \( H \) is Hurwitz, and thus the system \( \frac{dr}{d\tau} = Hr \) is globally exponentially stable. Therefore, there exists a Lyapunov function \( V(r, \tau) = r^TP(\tau)r \) that satisfies (9.3)–(9.5) in Khalil[45]. Furthermore, we have that \( \| C(\tau)h + D(\tau) \| \leq \| C(\tau) \| \| h \| + \| D(\tau) \| \). We note that \( \| D(\tau) \| \) is bounded as the functions \( g_2(b_1, x_0, \psi_{x_0}, 0) \) and \( g_3(b_1, x_0, \psi_{x_0}, 0) \) are continuous in \( b_1(\tau) \) from Assumptions 4.1 and 4.2, and \( \dot{b}_1(\tau) \) is bounded due to the asymptotic stability of the equilibrium point \( \dot{b}_1 = 0 \). Furthermore, since \( \lim_{\tau \to \infty} \dot{b}_1(\tau) = 0 \), it follows from the definitions of the functions \( g_2 \) and \( g_3 \) that \( \| C(\tau) \| \to 0 \) and \( \| D(\tau) \| \to 0 \). Therefore, from Lemma 9.5, 9.6 and Lemma 9.4 in
Khalil[45], we have that \( \lim_{r \to \infty} r(\tau) = 0 \) for all \( r(0) \in \mathbb{R}^3 \). Thus, it follows that the region of attraction for the system (B.29)-(B.31) is given by \( R_{b1} \times \mathbb{R}^3 \) where \( R_{b1} \) is the region of attraction of the equilibrium point \( z_0 - \gamma_1(x_0,0) \).

Next, we consider the remaining assumptions of the Tikhonov’s theorem. We have that functions \( f_x, f_z, S_x, S_z, F_x, F_z, \sigma_x\sigma_x^T, \sigma_x[\sigma_x0]^T \) and \( \sigma_x\sigma_z^T \) and their first partial derivatives are continuously differentiable from Assumptions 4.1 and 4.2. Under Assumption 4.1 we also have that the \( \frac{\partial f_1(x,z,t,0)}{\partial z}, \frac{\partial F_1(x,z,t,0)}{\partial z} \) have continuous first partial derivatives with respect to their arguments. From Assumptions 4.1 and 4.3 the first partial derivatives of \( \gamma_1(x,t), \gamma_2(x,t)E[\psi_x], \gamma_2(x,t)E[\psi_x\psi_x^T], \gamma_3(x,E[\psi_x\psi_x^T],t) \) with respect to their arguments are also continuous. Under Assumption 4.4 there exists a unique, bounded solution for the reduced system (4.13) for \( t \in [0,t_1] \). Furthermore, as the moment dynamics (4.23) - (4.24) are linear in the variables \( E[\tilde{\psi}_x], E[\tilde{\psi}_x\tilde{\psi}_x^T] \) there exists a unique, bounded solution to (4.23) - (4.24) for \( t \in [0,t_1] \). Thus, the assumptions of the Tikhonov’s theorem are satisfied and applying the theorem to the set of moment equations in (4.18)-(4.22) and (4.23)-(4.26) yields the result (4.27)-(4.33).

B.6 Proof for convergence of reduced and the original moments

Here, we prove that the moments of the reduced system (4.13)-(4.16) can provide a good approximation for the moments of the original variables \( v \) and \( \xi \) in the system (4.1)-(4.2). We only provide a complete proof for the variable \( v \) since the proof for \( E[\xi] \) and \( E[\xi^2] \) can be derived in a similar manner. From equation (4.1)-(4.2) we have that \( v \) and \( \xi \) represent the original variables, and from Claim 1 we have that

\[
v = T^{-1}[x^T,z^T]^T \quad \text{and} \quad \xi = T^{-1}[\psi_x^T,\psi_x^T]^T.
\]

Then, let \( \bar{v} \) and \( \bar{\xi} \) represent the species concentrations obtained using the reduced
system (4.13)-(4.16). Therefore, we have that

\[ \bar{v} = T^{-1}[x^T, \bar{z}^T]^T \text{ and } \bar{\xi} = T^{-1}[\bar{\psi}_x^T, \bar{\psi}_z^T]^T. \]

Then, computing the error between the moments of the original species concentrations and the moments obtained using the reduced system, we have

\[ \|v - \bar{v}\| = \left\| T^{-1} \begin{bmatrix} x \\ z \end{bmatrix} - T^{-1} \begin{bmatrix} \bar{x} \\ \bar{z} \end{bmatrix} \right\|, \]

\[ = \left\| T^{-1} \begin{bmatrix} x - \bar{x} \\ z - \bar{z} \end{bmatrix} \right\|. \]

Using the Cauchy-Schwarz inequality, we have that

\[ \|v - \bar{v}\| \leq \|T^{-1}\|_F \left\| \begin{bmatrix} x - \bar{x} \\ z - \bar{z} \end{bmatrix} \right\|, \]

and using the definition of the Euclidean norm we obtain

\[ \|v - \bar{v}\| \leq \|T^{-1}\| \sqrt{\|x - \bar{x}\|^2 + \|z - \bar{z}\|^2}. \]

From Theorem 4.1, we have that \(\|x - \bar{x}\| \leq c_1 \epsilon\) and \(\|z - \bar{z}\| \leq c_2 \epsilon\) for appropriate constants \(c_1\) and \(c_2\). Thus, from (B.32) it follows that

\[ \|v - \bar{v}\| \leq C \epsilon, \]

for an appropriate constant \(C\). Similar result holds for the variables \(E[\xi]\) and \(E[\xi^2]\).
Appendix C

Derivations and simulation results pertaining to Chapter 5

C.1 Derivation of the moment dynamics of the mixed variable

Here, we show the derivation on the moment dynamics of the variable $y$ using the slow and fast variable approximations. First, we derive the first and second moment dynamics of the variables $v$ and $c$ of the reduced system in (5.3)-(5.6) as

$$\frac{d\mathbb{E}[v]}{dt} = \frac{X}{k_{d1}} - (1 - R)\mathbb{E}[v], \quad \mathbb{E}[c] = \frac{\mathbb{E}[v]\beta_1\rho_{T0}}{\delta_1(p_T + k_{d2})}, \tag{C.1} \tag{C.2}$$

$$\frac{d\mathbb{E}[v^2]}{dt} = 2\frac{X}{k_{d1}}\mathbb{E}[v] - 2(1 - R)\mathbb{E}[v^2] + \frac{\delta_1X}{\beta_1\rho_{T0}\Omega} + \frac{\delta_1(1 - R)\mathbb{E}[v]}{\beta_1\rho_{T0}\Omega}, \quad \mathbb{E}[c^2] = \frac{\mathbb{E}[v^2]\beta_1^2\rho_{T0}^2}{\delta_1^2(p_T + k_{d2})^2} + \frac{\mathbb{E}[v]\beta_1\rho_{T0}k_{d2}}{\delta_1p_T\Omega(p_T + k_{d2})^2}, \tag{C.3} \tag{C.4}$$

$$\mathbb{E}[vc] = \frac{\mathbb{E}[v^2]\beta_1\rho_{T0}}{\delta_1(p_T + k_{d2})}. \tag{C.5}$$

Since $y = v - \frac{\rho_T\delta_1}{\beta_1\rho_{T0}}c$, we can write the dynamics for the first moment of $y$ as $\frac{d\mathbb{E}[y]}{dt} = \frac{d\mathbb{E}[v]}{dt} - \frac{\rho_T\delta_1}{\beta_1\rho_{T0}} \frac{d\mathbb{E}[c]}{dt}$. Then, using the chain rule we obtain $\frac{d\mathbb{E}[y]}{dt} = \left(1 - \frac{\rho_T\delta_1}{\beta_1\rho_{T0}} \frac{d\mathbb{E}[c]}{dt}\right) \frac{d\mathbb{E}[v]}{dt}$. 

159
and with \( \frac{dE[c]}{dt} = \frac{\beta_1 \rho_{0\alpha}}{\delta_1 (\rho_T + k_{d2})} \) from (C.2), we can write

\[
\frac{dE[y]}{dt} = \left(1 - \frac{\rho_T}{(\rho_T + k_{d2})}\right) \frac{dE[v]}{dt}
\]

(C.6)

Considering the dynamics for the second moment of \( y \), we have \( \frac{dE[y^2]}{dt} \). Substituting the fast variable approximation in (5.6) we obtain

\[
\frac{dE[y^2]}{dt} = \frac{dE \left[ \left( v - \frac{\rho_T}{(\rho_T + k_{d2})} + \sqrt{\frac{\rho_T k_{d2} \rho_{0\alpha}(\rho_T + k_{d2})}{\beta_1 \rho_{0\alpha}(\rho_T + k_{d2})}} N_2 \right)^2 \right]}{dt}
\]

Since \( N_2 \) is a normal random variable independent of \( y \) with \( E[N_2] = 0 \) and \( E[N_2^2] = 1 \), we have that

\[
\frac{dE[y^2]}{dt} = \frac{k_{d2}^2}{(\rho_T + k_{d2})^2} \frac{dE[v^2]}{dt} + \frac{\beta_1 \rho_{0\alpha}}{\beta_1 \rho_{0\alpha} (\rho_T + k_{d2})^2} \frac{dE[v]}{dt}
\]

(C.7)

Substituting in (C.6) and (C.7) the expressions for the dynamics of \( E[v] \) and \( E[v^2] \) given in (C.1) and (C.3), yields the first and second moment dynamics of \( y \) as

\[
\frac{dE[y]}{dt} = \left(1 - \frac{\rho_T}{(\rho_T + k_{d2})}\right) \left( \frac{X}{k_{d1}} - (1 - R)E[v] \right)
\]

(C.8)

\[
\frac{dE[y^2]}{dt} = \frac{k_{d2}^2}{(\rho_T + k_{d2})^2} \left( \frac{2X}{k_{d1}} E[v] - 2(1 - R)E[v^2] + \frac{\delta_1 X}{\beta_1 \rho_{0\alpha} \Omega} \right.
\]

\[
+ \left. \frac{\delta_1 (1 - R)E[v]}{\beta_1 \rho_{0\alpha} \Omega} \right) + \frac{\delta_1 \rho_{0\alpha}}{\beta_1 \rho_{0\alpha} (\rho_T + k_{d2})^2} \left( \frac{X}{k_{d1}} - (1 - R)E[v] \right).
\]

(C.9)

Next, in order to write the equations (C.8)–(C.9) in terms of \( E[y] \) and \( E[y^2] \), we use the moments of the fast variable approximation in (C.2) and (C.5) to express \( E[v] \) and \( E[v^2] \) in terms of \( E[y] \) and \( E[y^2] \). Towards this end, first consider \( E[v] = E[y] + \frac{\rho_T \delta_1}{\beta_1 \rho_{0\alpha}} E[c] \). Using the expression for \( E[c] \) from (C.2) we have

\[
E[v] = \frac{(k_{d2} + \rho_T)}{k_{d2}} E[y].
\]

(C.10)
Considering the second moments, we obtain

\[
E[v^2] = E\left[ (y + \frac{pr\delta_1}{\beta_1p_{T0}}c)^2 \right]
= E[y^2] + 2 \frac{pr\delta_1}{\beta_1p_{T0}} E[yc] + \frac{p^2r^2\delta^2_1}{\beta^2_1p^2_{T0}} E[c^2], \tag{C.11}
\]

\[
E[vc] = E\left[ (y + \frac{pr\delta_1}{\beta_1p_{T0}}c)c \right] = E[yc] + \frac{p\delta_1}{\beta_1p_{T0}} E[c^2]. \tag{C.12}
\]

Using \(E[yc] = E[vc] - \frac{p\delta_1}{\beta_1p_{T0}} E[c^2]\) from (C.12) in (C.11) we obtain \(E[v^2] = E[y^2] + 2 \frac{pr\delta_1}{\beta_1p_{T0}} E[vc] - \frac{p^2r^2\delta^2_1}{\beta^2_1p^2_{T0}} E[c^2]\). Then, substituting for \(E[vc]\) and \(E[c^2]\) from (C.4)–(C.5), yields \(E[v^2] = E[y^2] + 2 \frac{E[\sigma^2]pr}{(pr + k)} - \frac{E[\sigma^2]p^2}{(pr + k)} - \frac{E[\sigma\delta pr]}{\delta_1pr_0(pr + k)^2}\). Simplifying further and using the expression for \(E[v]\) from (C.10) yields

\[
E[v^2] = \frac{(kd_2 + pr)^2}{kd_2^2} E[y^2] - \frac{\delta_1(kd_2 + pr)prE[y]}{kd_2^2 \Omega \beta_1p_{T0}}. \tag{C.13}
\]

Then, substituting (C.10) and (C.13) in (C.8)–(C.9) and simplifying further yields the moment dynamics of the mixed variable given in (5.15)–(5.16).

### C.2 Derivations for Section 5.3

**Noise error of protein G**

Evaluating the magnitude of the frequency response, we find that the noise error is given by equation (5.21) where the function \(A(p_T, \tilde{\omega})\) is in the form \(A(p_T, \tilde{\omega}) = \frac{(1-R)^2(\tilde{\omega}^2)}{(\beta_1p_{T0}/k_d)^2\beta_2^2\delta_4^2p_T^2D_1D_2}\), with

\[
N_1 = -2\delta_4^2(R - 1)\tilde{\omega}^2(k_{d2} + p_T)(2\beta_2p_T + \delta_2(k_{d2} + p_T)),
\]

\[
N_2 = \delta_1^2(4\delta_2^2p_T^2\tilde{\omega}^2 + 8\beta_2\delta_2p_T\tilde{\omega}^2(k_{d2} + p_T))
+ \delta_1^2(\delta_2^2(k_{d2} + p_T)^2(4R^2 - 8R + 5\tilde{\omega}^2 + 4)),
\]

\[
N_3 = -8\delta_1\delta_2(R - 1)(k_{d2} + p_T)(\beta_2p_T + \delta_2(k_{d2} + p_T)),
\]

\[
N_4 = 4\delta_2^2(\beta_2p_T + \delta_2(k_{d2} + p_T))^2
+ \delta_1^4\omega^2(k_{d2} + p_T)^2((1 - R)^2 + \tilde{\omega}^2),
\]

161
To identify the change in $A(pt, \omega)$ with $pt$, we consider the derivative of $A(pt, \omega)$ with respect to $pt$. Evaluating the derivative at $\omega = 0$, yields $\frac{\partial A(pt, \omega)}{\partial pt} = \frac{2k^3}{(\beta_1 pt/ko) \beta_2^2 k_2^2 p_t^2 (\delta_1 k_2 + \delta_2 (k_2 + pt))}$, where

\[
N_{d1} = \delta_2 (\beta_2^2 pt^3 + \beta_2 \delta_2 pt (k_2^2 + 3k_2^2 pt + 2pt^2) + \delta_2^2 (k_2^2 + pt)^3),
\]
\[
N_{d2} = \delta_1^2 k_2^2 (\beta_2 pt + 3\delta_2 (k_2 + pt)),
\]
\[
N_{d3} = \delta_1 \delta_2 k_2 (\beta_2 pt (2k_2 + 3pt) + 3\delta_2 (k_2 + pt)^2).
\]

It follows that the derivative is negative for all parameter conditions. The function $A(pt, \omega)$ is a rational polynomial function in $pt$ and is continuous with respect to $\omega$. Thus, we $\frac{\partial A(pt, \omega)}{\partial pt}$ is continuous with respect to $\omega$ and will remain negative in a neighborhood of $\omega = 0$. Therefore, the function $A(pt, \omega)$ is decreasing with $pt$ for sufficiently small $\omega$.

C.3 Derivations for Section 5.6

Noise error of protein G

Using the set of moment equations (5.71) - (5.77), we can calculate the coefficient of variation of the protein G as

$$\sqrt{\frac{\text{E}[\tilde{q}_g^2]}{q}} = \sqrt{A(Y)},$$

where we have

$$A(Y) = \frac{\delta ((k_1 Z + k_2 Y)N_{g1} + (ak_1 Y Z + k_d k_1 Z + k_d k_2 Y)N_{g2})}{a \beta k_1 pt Y Z (N_{g2} + (k_1 Z + k_2 Y)D_{g1})}$$

with
Taking the derivative of $A(Y)$ with respect to $Y$, yields

$$\frac{\delta k_d(k_1 Z(N_{g2} + N_{gd1})^2 + \beta k_d k_2 p_T Y^2 (k_1 Z + k_2 Y) N_{g2} N_{g2} N_{g3})}{a \beta k_1 p_T Y^2 Z (N_{g2} + N_{gd1})^2}$$

where

$$N_{gd1} = (k_1 Z + k_2 Y)(a^2 k_1^2 Y^2 Z^2 + k_d^2 (k_1 Z + k_2 Y)^2$$
$$+ k_d k_1 Z(2a Y + p_T)(k_1 Z + k_2 Y),$$

$$N_{gd2} = a k_1 Z(2k_1 Z(\delta + k_1 Z) + 3k_1 k_2 Y Z + k_2^2 Y^2)$$
$$+ k_d k_2 (k_1 Z + k_2 Y)^2,$$

$$N_{gd3} = (a k_1 Y Z + k_d k_1 Z + k_d k_2 Y).$$

Since all the parameter values are positive, it can be seen that the derivative of the function $A(Y)$ with respect to $Y$ is negative.

**Noise error of protein $X^*$**

Similarly, the coefficient of variation of protein $X^*$ can be found by using the set of moment equations (5.71) - (5.77) and the expression (5.78), which yields

$$\sqrt{\mathbb{E}[\tilde{v}_2^2]} = \sqrt{B(Y)}$$

where
\[ B(Y) = \frac{a^2 k_1^2 k_1 Y^3 Z^2 + k_d (k_1 Z + k_1 Y) N_{x1} + k_d^2 k_1 Y (k_1 Z + k_1 Y)^2}{a k_1 Y Z D_{x1}} \]

with

\[ N_{x1} = 2 k_1 k_2 Y Z (a Y + p_T) + k_1^2 p_T Z^2 + k_2^3 p_T Y^2, \]
\[ D_{x1} = a^2 k_1^2 Y^2 Z^2 + k_d (k_1 Z + k_2 Y) (k_1 Z (2 a Y + p_T) + k_2 p_T Y) + k_d^2 (k_1 Z + k_2 Y)^2. \]

Calculating the derivative of \( B(Y) \) with respect to \( Y \), yields

\[-\frac{k_d p_T (k_1 Z + k_2 Y) (a^2 k_1^2 Y^2 Z^2 (3 k_1 Z + k_2 Y) + N_{x2d})}{a Y^2 D_{x1}^2} \]

where

\[ N_{x1d} = (k_1^2 Z^2 (4 a Y + p_T) + 2 k_1 k_2 Y Z (a Y + p_T) + k_2^3 p_T Y^2), \]
\[ N_{x2d} = k_d (k_1 Z + k_2 Y) N_{x1d} + k_d^2 (k_1 Z + k_2 Y)^3. \]

We have that all system parameters are positive, and therefore, it follows that the derivative of \( B(Y) \) with respect to \( Y \) is negative for all parameter values.

C.4 Simulations of the full and reduced systems

C.4.1 Simulation results for Section 5.5

Figure C-1 shows the error in the moments between the dynamics of the full system (5.32)-(5.35) and the reduced system (5.37)-(5.40). The simulations are performed using the Euler-Maruyama method for stochastic differential equations and the moments are calculated using 100,000 simulation runs.

This system does not satisfy the sufficient conditions for existence of a unique solution in [16], and we note that the existence of a unique, well-defined solution for chemical Langevin equations is an ongoing research question [91, 74]. Therefore, in this case, we choose parameter conditions that give sufficiently large molecular counts,
Figure C-1: Errors in the first and second order moments. Average of 100,000 runs. The parameters used are $k = 10$, $\delta_1 = 0.1$, $\delta_1 = 1$, $k_{d1} = 1000$, $k_{d2} = 1000$, $P_{t1} = 1000$, $P_{t2} = 1000$, $\beta = 1$, $y(0) = 70$, $G(0) = 60$, $C_1(0) = 20$ and $C_2(0) = 20$.

in order to increase the probability that the argument of the square-root term remains positive. In this example, the argument of the square-root remained positive for all the simulations performed and used to numerically determine the sample means.
C.4.2 Simulation results for Section 5.6

Figure C-2 illustrates simulation results for the second moments of the variables $\psi_v$, $\psi_g$ and $\psi_c$ in the full system (5.47)-(5.49) and reduced system (5.53)-(5.58). The simulations were performed using the Euler-Maruyama method for the stochastic differential equations and the moments are computed using the average of 300,000 simulation runs. We use zero initial conditions for all the variables in both the original and the reduced systems, and thus the first moments of the stochastic fluctuations remain zero at all times. For the second moments, as shown in Figure 2, it can be seen that, as $\epsilon$ tends to zero the moments of the original system tends to the moments of the reduced system. For the variables $\psi_v$ and $\psi_c$, there is a good agreement between the moments of the original and reduced-order systems at $\epsilon = 0.1$. For $\psi_g$, further reducing $\epsilon$ to 0.01 provides a better approximation of the second moment.

Figure C-2: Moments of the original and reduced systems. (a) Second moments of $\psi_v$ and $\overline{\psi}_v$. (b) Second moments of $\psi_g$ and $\overline{\psi}_g$. (c) Second moments of $\psi_c$ and $\overline{\psi}_c$. 

166
Appendix D

Derivations pertaining to Chapter 6

D.1 Stability of the slow manifold

\[ \varepsilon c_{ij} = \frac{\delta_{ij} \beta_{ij}}{\beta k_{dij}} (x_j - \sum_{p \in \mathcal{Y}_j} c_{pj})^n (p_{T_i} - c_{ij}) - \delta_j \frac{\beta_{ij}}{\beta} c_{ij} \]  \hspace{1cm} (D.1)

Note that for a given \( c_{ij} \) the dynamics are dependent only on the \( c_{pj} \) where \( c_{pj} \in \sum_{\forall p \in \mathcal{Y}_j} \). Therefore, let \( < j_1, \ldots, j_{|\mathcal{Y}_j|} > \) be an ordered list of the node indexes of the proteins in \( \mathcal{Y}_j \). Then, with \( c_j = [c_{j_1}; \ldots; c_{j_{|\mathcal{Y}_j|}}] \), we have

\[ \varepsilon \dot{c}_j = \mathcal{G}_j(x, c_j). \]  \hspace{1cm} (D.2)

Then to determine the stability of \( c_j \) we take

\[ \frac{\partial \mathcal{G}_j(x, c_j)}{\partial c_j} = - \left[ (\delta_j \frac{\beta_{j1}}{\beta} + \frac{\delta_j \beta_{j1}}{\beta k_{d_{j1}}} (x_j - \sum_{p \in \mathcal{Y}_j} c_{pj})^n) + n \frac{\delta_j \beta_{j1}}{\beta k_{d_{j1}}} (p_{T_{j1}} - c_{j1}) y_j^{n-1} \right. \]

\[ + \ldots \]

\[ \left. \ldots \right] \]

\[ + \left[ n \frac{\delta_{j_{|\mathcal{Y}_j|}} \beta_{j_{|\mathcal{Y}_j|}}}{\beta k_{d_{j_{|\mathcal{Y}_j|}}}} (p_{T_{j_{|\mathcal{Y}_j|}}} - c_{j_{|\mathcal{Y}_j|}}) y_j^{n-1} \right. \]

\[ + \ldots \]

\[ \left. \ldots \right] \]

\[ + \left( \delta_j \frac{\beta_{j_{|\mathcal{Y}_j|}}}{\beta} + \frac{\delta_j \beta_{j_{|\mathcal{Y}_j|}}}{\beta k_{d_{j_{|\mathcal{Y}_j|}}}} y_j^n \right) + n \frac{\delta_{j_{|\mathcal{Y}_j|}} \beta_{j_{|\mathcal{Y}_j|}}}{\beta k_{d_{j_{|\mathcal{Y}_j|}}}} (p_{T_{j_{|\mathcal{Y}_j|}}} - c_{j_{|\mathcal{Y}_j|}}) y_j^{n-1} \right] \]
which can then be expressed as

$$
\frac{\partial \hat{g}_j(x_j, c_j)}{\partial c_j} = - \left[ \begin{array}{ccc}
\frac{(PT_{j1}/k_{d1j})y_j^{n-1}}{n(y_j^n/k_{d1j}+1)} & 0 & 0 \\
0 & \ddots & 0 \\
0 & 0 & \frac{(PT_{jn}/k_{d1j})y_j^{n-1}}{n(y_j^n/k_{d1j}+1)}
\end{array} \right]
\left[ \begin{array}{c}
\frac{(1+y_j^n)}{n(y_j^n/k_{d1j}+1)} \\
0 & \ddots & 0 \\
0 & \frac{1}{y_j^n/k_{d1j}+1}
\end{array} \right]
\begin{bmatrix}
1 & \cdots & 1 \\
1 & \ddots & 1 \\
1 & \cdots & 1
\end{bmatrix}
$$

$$= - A(D + J).$$

We have that $D$ is a positive-definite matrix and $J$ is a positive-semidefinite matrix, which implies $D + J$ is positive-definite. The matrix $A$ is also positive definite and we have that, $Q = A(D + J)$ is similar to $A^{-1/2}A(D + J)A^{1/2} = A^{1/2}(D + J)A^{1/2}$. We have that $A^{1/2}(D + J)A^{1/2}$ is congruent to $(D + J)$, which implies that the $A(D + J)$ have all positive eigenvalues. We therefore have that $-A(D + J)$ is a Hurwitz matrix.

### D.2 Matrix inverse

**Claim D.1.** Consider the matrix

$$M = \begin{bmatrix}
a_1 & 1 & \cdots & 1 \\
1 & a_2 & \cdots & 1 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \cdots & a_n
\end{bmatrix}$$
with $a_i \neq 1$. We have that

$$M^{-1} = \begin{bmatrix}
\frac{1}{(a_1-1)} & 0 & \ldots & 0 \\
0 & \frac{1}{(a_2-1)} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \frac{1}{(a_n-1)}
\end{bmatrix}
- \left(1 + \sum_{i=1}^{n} \frac{1}{(a_i-1)}\right)^{-1}
\begin{bmatrix}
\frac{1}{(a_1-1)} & \frac{1}{(a_2-1)} & \frac{1}{(a_3-1)} & \ldots & \frac{1}{(a_n-1)} \\
\frac{1}{(a_2-1)} & \frac{1}{(a_3-1)} & \frac{1}{(a_4-1)} & \ldots & \frac{1}{(a_n-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{(a_n-1)} & \frac{1}{(a_{n-1}-1)} & \frac{1}{(a_{n-2}-1)} & \ldots & \frac{1}{(a_{n-1}-1)}
\end{bmatrix}
$$

Proof. Note that we can write the matrix $M$ in the form

$$M = \begin{bmatrix}
a_1 & 1 & \ldots & 1 \\
1 & a_2 & \ldots & 1 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \ldots & a_n
\end{bmatrix}^{-1}
\begin{bmatrix}
a_1 - 1 & 0 & \ldots & 0 \\
0 & a_2 - 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & a_n - 1
\end{bmatrix}
+ \begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix}.$$

Then taking

$$A = \begin{bmatrix}
a_1 - 1 & 0 & \ldots & 0 \\
0 & a_2 - 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & a_n - 1
\end{bmatrix}$$
we can apply the Woodbury matrix formula which yields

\[
M^{-1} = A^{-1} - A^{-1} U (C^{-1} + VA^{-1}U)^{-1} VA^{-1},
\]

\[
= \begin{bmatrix}
1/(a_1 - 1) & 0 & \ldots & 0 \\
0 & 1/(a_2 - 1) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1/(a_n - 1)
\end{bmatrix}
\begin{bmatrix}
1 \\
1/(a_2 - 1) \\
\vdots \\
1/(a_n - 1)
\end{bmatrix}
\]

\[
\left(1 + \begin{bmatrix}
1 & 1 & \ldots & 1
\end{bmatrix}
\begin{bmatrix}
1/(a_1 - 1) & 0 & \ldots & 0 \\
0 & 1/(a_2 - 1) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1/(a_n - 1)
\end{bmatrix}
\begin{bmatrix}
1 \\
1/(a_2 - 1) \\
\vdots \\
1/(a_n - 1)
\end{bmatrix}
\right)^{-1}
\]

\[
\begin{bmatrix}
1/(a_1 - 1) & 0 & \ldots & 0 \\
0 & 1/(a_2 - 1) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1/(a_n - 1)
\end{bmatrix}
\begin{bmatrix}
1 \\
1/(a_2 - 1) \\
\vdots \\
1/(a_n - 1)
\end{bmatrix}
\]
D.3 System models for incoherent feedforward loop \(^1\)

The chemical reactions for the network in Figure 6-2 can be written as

\[
Y_1 + D_{20} \xrightarrow{\alpha_{31}} D_{21}, \\
Y_1 + D_{30} \xrightarrow{\alpha_{31}} D_{31}, \\
Y_2 + D_{30} \xrightarrow{\alpha_{32}} D_{32}, \\
Y_1 + D_{32} \xrightarrow{\alpha_{31}} D_{33}, \\
Y_2 + D_{31} \xrightarrow{\alpha_{32}} D_{33}, \\
D_{10} \xrightarrow{\pi_{10}} Y_1 + D_{10}, \\
D_{21} \xrightarrow{\pi_{21}} Y_2 + D_{21}, \\
D_{31} \xrightarrow{\pi_{31}} Y_3 + D_{31},
\]

in which \( D_{21} \) is the complex formed by the binding of \( Y_1 \) to the promoter site \( D_{20} \) in node 2 and \( D_{31}, D_{32}, D_{33} \) are the complexes formed due to the bindings of inputs \( Y_1 \) and \( Y_2 \) in node 3 and we have assumed that the binding/unbinding rates between \( Y_1 \) and the promoter sites of node 2 and 3 are equal. We also have the conservation laws

\[
P_{T2} = D_{20} + D_{21} \text{ and } P_{T3} = D_{30} + D_{31} + D_{32} + D_{33}.
\]

In order to derive the LNA models based on stochastic QSSA and tQSSA\(^+\), we first compute the macroscopic reaction rates as

\[
\frac{dy_1}{dt} = \beta_1 p_{T1} - \delta_1 y_1 - \alpha_{31} y_1 (p_{T2} - d_{21}) + \beta_{31} d_{21} - \alpha_{31} y_1 (p_{T3} - d_{31} - d_{32} - d_{33})
\]

\(^1\)The model derivations in this section were performed by Simone Bruno.
\[ \frac{dy_2}{dt} = \beta_2 d_{21} - \delta_2 y_2 - \alpha_{32} y_2 (p_{T3} - d_{31} - d_{32} - d_{33}) + \beta_{32} d_{32} - \alpha_{32} y_2 d_{31} + \beta_{32} d_{33}, \]
\[ \frac{dy_3}{dt} = \beta_3 d_{31} - \delta_3 y_3, \]
\[ \frac{dd_{21}}{dt} = \alpha_{31} y_1 (p_{T2} - d_{21}) - \beta_{31} d_{21}, \]
\[ \frac{dd_{31}}{dt} = \alpha_{31} y_1 (p_{T3} - d_{31} - d_{32} - d_{33}) - \beta_{31} d_{31} - \alpha_{32} y_2 d_{31} + \beta_{32} d_{33}, \]
\[ \frac{dd_{32}}{dt} = \alpha_{32} y_2 (p_{T3} - d_{31} - d_{32} - d_{33}) - \beta_{32} d_{32} - \alpha_{31} y_1 d_{32} + \beta_{31} d_{33}, \]
\[ \frac{dd_{33}}{dt} = \alpha_{31} y_1 d_{32} + \alpha_{32} y_2 d_{31} - (\beta_{31} + \beta_{32}) d_{33}. \]

We have that the binding reactions are much faster than protein production/decay and thus we can write \( \epsilon = \delta_1 / \beta_{31} \). Then, defining the dissociation constants \( k_{d31} = \beta_{31} / \alpha_{31}, \) \( k_{d32} = \beta_{32} / \alpha_{32}, \) and the ratios \( a = \beta_{31} / \beta_{31}, \) \( b = \beta_{32} / \beta_{31}, \) we can write

\[ \frac{dy_1}{dt} = \beta_1 p_{T1} - \delta_1 y_1 - \alpha_{31} y_1 (p_{T2} - d_{21}) + \beta_{31} d_{21} - \alpha_{31} y_1 (p_{T3} - d_{31} - d_{32} - d_{33}) \]
\[ + \beta_{31} d_{31} - \alpha_{31} y_1 d_{32} + \beta_{31} d_{32}, \]
\[ \frac{dy_2}{dt} = \beta_2 d_{21} - \delta_2 y_2 - \alpha_{32} y_2 (p_{T3} - d_{31} - d_{32} - d_{33}) + \beta_{32} d_{32} - \alpha_{32} y_2 d_{31} + \beta_{32} d_{33}, \]
\[ \frac{dy_3}{dt} = \beta_3 d_{31} - \delta_3 y_3, \]
\[ \epsilon \frac{dd_{21}}{dt} = (\delta_1 / k_{d31}) y_1 (p_{T2} - d_{21}) - \delta_1 d_{21}, \]
\[ \epsilon \frac{dd_{31}}{dt} = a(\delta_1 / k_{d31}) y_1 (p_{T3} - d_{31} - d_{32} - d_{33}) - a\delta_1 d_{31} - b(\delta_1 / k_{d32}) y_2 d_{31} + b\delta_1 d_{33}, \]
\[ \epsilon \frac{dd_{32}}{dt} = b(\delta_1 / k_{d32}) y_2 (p_{T3} - d_{31} - d_{32} - d_{33}) - b\delta_1 d_{32} - a(\delta_1 / k_{d31}) y_1 d_{32} + a\delta_1 d_{33}, \]
\[ \epsilon \frac{dd_{33}}{dt} = a(\delta_1 / k_{d31}) y_1 d_{32} + b(\delta_1 / k_{d32}) y_2 d_{31} - (a\delta_1 + b\delta_1) d_{33}. \]

### D.3.1 Hill function model

We next derive the hill function model by taking the quasi-steady state of the promoter-TF complexes \( d_{21}, d_{31}, d_{32}, d_{33} \), which yields

\[ \frac{d\hat{y}_1}{dt} = \pi_{10} p_{T1} - \delta_1 \hat{y}_1, \]
\[
\begin{align*}
\frac{d\tilde{y}_2}{dt} &= \frac{\pi_{21}}{k_{d31} + \tilde{y}_1} \frac{p_{T2} \tilde{y}_1}{k_{d31}} - \delta_2 \tilde{y}_2, \\
\frac{d\tilde{y}_3}{dt} &= \frac{\pi_{31}}{1 + \tilde{y}_1/k_{d31} + \tilde{y}_2/k_{d32} + \tilde{y}_1 \tilde{y}_2/k_{d31} k_{d32}} \frac{p_{T3} \tilde{y}_1/k_{d31}}{k_{d31} k_{d32}} - \delta_3 \tilde{y}_3, \\
\frac{d\tilde{\xi}_1}{dt} &= -\delta_1 \tilde{\xi}_1 + \sqrt{\pi_{10} p_{T1} \Gamma_1} - \sqrt{\delta_1} \tilde{y}_1 \Gamma_2, \\
\frac{d\tilde{\xi}_2}{dt} &= -\delta_2 \tilde{\xi}_2 + \pi_{21} \frac{p_{T2} \tilde{y}_1}{k_{d31} + \tilde{y}_1} \frac{\Gamma_3}{\sqrt{\delta_2} \tilde{y}_2 \Gamma_4}, \\
\frac{d\tilde{\xi}_3}{dt} &= -\delta_3 \tilde{\xi}_3 + \pi_{31} \frac{p_{T3} k_{d31} k_{d32} (k_{d32} + \tilde{y}_2)}{y_1 k_{d32} + \tilde{y}_2 k_{d31} + \tilde{y}_1 \tilde{y}_2 + k_{d31} k_{d32}} \frac{\Gamma_3}{\sqrt{\delta_2} \tilde{y}_2 \Gamma_4} + \frac{\pi_{31} p_{T3} k_{d33} f_1 (k_{d31} + \tilde{y}_1)}{\tilde{y}_1 k_{d32} + \tilde{y}_2 k_{d31} + \tilde{y}_1 \tilde{y}_2 + k_{d31} k_{d32}} \frac{\tilde{\xi}_2}{\sqrt{\delta_3} \tilde{y}_3 \Gamma_5} - \sqrt{\delta_3} \tilde{y}_3 \Gamma_6,
\end{align*}
\]

where \( \xi_i \) denotes the stochastic fluctuations around each \( y_i \).

### D.3.2 Stochastic tQSSA+ model

In order to derive the stochastic tQSSA+ model we first need to identify the slow and fast variables in the system. Therefore, via a coordinate transformation we define the slow variables \( x_1 = y_1 + d_{21} + d_{31} + d_{33} = y_1 + d_{21} + c_{31}, \ x_2 = y_2 + d_{32} + d_{33} = y_2 + c_{32}, \ x_3 = y_3 \) and the fast variables \( d_{21}, \ c_{31}, \ c_{32}, \ d_{31} \).

Then, taking \( \epsilon = 0 \), we can write the reduced order model for the slow variables \( x_1, x_2, x_3 \) and the stochastic fluctuations around these variables denoted by \( \chi_1, \chi_2, \chi_3 \) as

\[
\begin{align*}
\frac{dx_1}{dt} &= \pi_{10} p_{T1} - \delta_1 (x_1 - \tilde{d}_{21} - \tilde{c}_{31}), \\
\frac{dx_2}{dt} &= \pi_{21} \tilde{d}_{21} - \delta_2 (x_2 - \tilde{c}_{32}), \\
\frac{dx_3}{dt} &= \pi_{31} \tilde{d}_{31} - \delta_3 x_3, \\
\frac{d\chi_1}{dt} &= -\delta_1 \chi_1 + \delta_1 \tilde{\psi}_{d_{21}} + \delta_1 \tilde{\psi}_{c_{31}} + \sqrt{\pi_{10} p_{T1} \Gamma_1} - \sqrt{\delta_1} (x_1 - \tilde{d}_{21} - \tilde{c}_{31}) \Gamma_2, \\
\frac{d\chi_2}{dt} &= -\delta_2 \chi_2 + \delta_2 \tilde{\psi}_{c_{32}} + \pi_{21} \tilde{\psi}_{d_{21}} + \sqrt{\pi_{21} \tilde{d}_{21} \Gamma_3} - \sqrt{\delta_2} (x_2 - \tilde{c}_{32}) \Gamma_4, \\
\frac{d\chi_2}{dt} &= -\delta_3 \chi_3 + \pi_{31} \tilde{\psi}_{d_{31}} + \sqrt{\pi_{31} \tilde{d}_{31} \Gamma_5} - \sqrt{\delta_3} x_3 \Gamma_6,
\end{align*}
\]
where

\[ d_{21} = \frac{x_1 + p_{T2} + k_{d31} - \bar{c}_{31}}{2} + \sqrt{\left(\frac{x_1 + p_{T2} + k_{d31} - \bar{c}_{31}}{2}\right)^2 - 4p_{T2}(x_1 - c_{31})}, \]  

(D.3)

\[ c_{31} = \frac{x_1 + p_{T3} + k_{d31} - d_{21}}{2} + \sqrt{\left(\frac{x_1 + p_{T3} + d_{21}}{2}\right)^2 - 4p_{T3}(x_1 - d_{21})}, \]  

(D.4)

\[ c_{32} = \frac{x_2 + p_{T3} + k_{d32}}{2} + \sqrt{\left(\frac{x_2 + p_{T3} + k_{d32}}{2}\right)^2 - 4p_{T3}x_2}, \]  

(D.5)

\[ \dd{3} = \frac{k_{d31}(x_1 - d_{21} - \bar{c}_{31})\bar{c}_{32} + k_{d32}(x_2 - \bar{c}_{32})\bar{c}_{31}}{k_{d31}(x_1 - d_{21} - \bar{c}_{31}) + k_{d32}(x_2 - \bar{c}_{32}) + 2k_{d31}k_{d32}}, \]  

(D.6)

\[ \dd{d3} = \frac{(p_{T2} - d_{21})(p_{T3} + x_1 - d_{21} - 2\bar{c}_{31} + k_{d31}) - (p_{T3} - \bar{c}_{31})x_1}{(p_{T2} + x_1 - \bar{c}_{31} - 2d_{21} + k_{d31})(p_{T3} + x_1 - d_{21} - 2\bar{c}_{31} + k_{d31}) - (p_{T2} - d_{21})(p_{T3} - \bar{c}_{31})}, \]  

(D.7)

\[ \dd{c31} = \frac{(p_{T3} - \bar{c}_{31})(x_1 - \dd{d3})}{p_{T3} + x_1 - d_{21} - 2\bar{c}_{31} + k_{d31}}, \]  

(D.8)

\[ \dd{c32} = \frac{(p_{T3} - \bar{c}_{32})(x_2)}{p_{T3} + x_2 - 2\bar{c}_{32} + k_{d32}}, \]  

(D.9)

\[ \dd{d3} = \frac{k_{d31}(\bar{c}_{32} - \dd{d3})x_1 + k_{d32}(\bar{c}_{31} - \dd{d3})x_2 - k_{d31}(\bar{c}_{32} - \dd{d3})\dd{d3}}{k_{d31}(x_1 - d_{21} - \bar{c}_{31}) + k_{d32}(x_2 - \bar{c}_{32}) + 2k_{d31}k_{d32}} + \frac{(-k_{d31}\bar{c}_{32} + k_{d32}(x_2 - \bar{c}_{32}) + k_{d31}d_{33})\dd{c31} + (k_{d31}(x_1 - d_{21} - \bar{c}_{31}) - k_{d32}\bar{c}_{31} + k_{d32}d_{33})\dd{c32}}{k_{d31}(x_1 - d_{21} - \bar{c}_{31}) + k_{d32}(x_2 - \bar{c}_{32}) + 2k_{d31}k_{d32}}, \]  

(D.10)

Next, considering the approximations for the fast variables given by the complexes, we have that approximations for the deterministic counterparts of the fast variables are given by (D.3)-(D.6). The approximations for the stochastic fluctuations can be found by considering a Lyapunov equation of the form (4.17) in Section 4 and we obtain that the approximations \( \Upsilon_{21}, \zeta_{31}, \zeta_{32} \) corresponding to \( d_{21}, c_{31}, c_{32} \) are given by

\[ \Upsilon_{21} = \dd{d3}, \]  

\[ \zeta_{31} = \dd{c31}, \]  

\[ \zeta_{32} = \dd{c32}, \]  

\[ \zeta_{32} = \dd{c32}, \]
where $\psi_{d_{21}}, \psi_{c_{31}}, \psi_{c_{32}}$ are in (D.7)-(D.10) and

\[
\begin{align*}
g_{d_{21}} &= \frac{1}{\sqrt{S + 2\sqrt{D}}}((h_1 + \sqrt{D})N_1 + h_2 N_2), \\
g_{c_{31}} &= \frac{1}{\sqrt{S + 2\sqrt{D}}}(h_3 N_1 + (h_4 + \sqrt{D})N_2), \\
g_{c_{32}} &= \sqrt{\frac{(x_2 - c_{32})(p_{T3} - c_{32} + k_{d32}c_{32})}{2((p_{T3} - x_2 - 2c_{32}) - 2c_{32} + k_{d32})}}N_3,
\end{align*}
\]

where $S = h_1 + h_2$, $D = h_1h_4 - h_2h_3$ with

\[
\begin{align*}
h_1 &= \frac{-a(d + g) + 2def(a - b)}{d(d + g)(1 - 2ef)}, \\
h_2 &= \frac{b(ga + db)}{(d + g)(1 - 2ef)}, \\
h_3 &= \frac{f(ga + db)}{(d + g)(1 - 2ef)}, \\
h_4 &= \frac{-b(d + g) + 2efg(b - a)}{g(d + g)(1 - 2ef)}.
\end{align*}
\]

and

\[
\begin{align*}
a &= (x_1 - \tilde{d}_{21} - \tilde{c}_{31})(p_{T2} - \tilde{d}_{21}) - k_{d31}\tilde{d}_{21}, \\
b &= (x_1 - \tilde{d}_{21} - \tilde{c}_{31})(p_{T3} - \tilde{c}_{31}) - k_{d31}\tilde{c}_{31}, \\
d &= (2\tilde{d}_{21} - p_{T2} + \tilde{c}_{31} - x_1 - k_{d31}), \\
e &= (p_{T2} - \tilde{d}_{21}), \\
f &= (2\tilde{c}_{31} - p_{T3} + \tilde{d}_{21} - x_1 - k_{d31}), \\
g &= (p_{T3} - \tilde{c}_{31}).
\end{align*}
\]

Then, simulating these system dynamics and considering the change of variables to the original species concentrations given by $y_1 = x_1 - d_{21} - c_{31}$, $y_2 = x_2 - c_{32}$, $y_3 = x_3$ and $\xi_1 = \chi_1 - \zeta_{21} - \zeta_{31}$, $\xi_2 = \chi_2 - \zeta_{32}$, $\xi_3 = \chi_3$ we obtain the moment dynamics in Figures 6-3 and 6-4.
Bibliography


