Analytical and Computational Approaches to Nonlinear Dynamics in Enzyme Kinetics and Sleep-Wake Regulation

by

Kashvi Srivastava

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied and Interdisciplinary Mathematics and Scientific Computing)
in the University of Michigan
2025

Doctoral Committee:

Professor Victoria Booth, Co-Chair
Professor Santiago Schnell, Co-Chair, Dartmouth College
Professor Daniel Beard
Dr. Justin Eilertsen, Mathematical Reviews
Professor Peter Miller

Kashvi Srivastava
kashvi@umich.edu
ORCID iD: 0000-0003-3401-4032

© Kashvi Srivastava 2025

DEDICATION

This thesis is dedicated to my sister, Srushti, who has inspired me in so many ways. Thank you for being the love, light and laughter in my life.

ACKNOWLEDGEMENTS

I would like to thank the Department of Mathematics at the University of Michigan, the College of Science at the University of Notre Dame and the National Center for Complementary and Integrative Health at the National Institutes of Health for providing me with financial support during my Ph.D. I am also grateful to the Michigan Center for Applied and Interdisciplinary Mathematics, Michigan Institute for Computational Discovery and Engineering and the Rackham Graduate School at University of Michigan for partially funding my Ph.D.

I would like to express heartfelt gratitude to my advisor, Professor Victoria Booth, for her unwavering support at every stage of my graduate studies. She has always provided me her generous time and guidance, both professional and personal, and helped me grow immensely over the past few years. She has been extremely kind, patient and encouraging and created a safe space for me to turn to in challenging times. I couldn't have asked for a better advisor and I will always cherish our discussions and memories together.

I would also like to thank my advisor, Professor Santiago Schnell, for his advice and guidance on research and writing. He has inspired me with his dedication and academic rigor and pushed me to pursue excellence. I am grateful to him for welcoming me to his lab and helping me secure different milestones during my Ph.D. and beyond.

I would especially like to thank my mentor, collaborator, and dear friend, Dr. Justin Eilertsen. A lot of what I have learnt about enzyme kinetics and scientific research is because of him. I extend my deepest gratitude to him for all the time and effort he has put towards my Ph.D., as well as for being a sanity check and a motivating voice for me. I couldn't have done this without him and I am extremely lucky to have worked with him.

I would like to thank our collaborator, Professor Cecilia Diniz Behn, for her mentorship and guidance. I have truly enjoyed working with her over the last few years and being encouraged by her kind and thoughtful discussions. I would like to thank my committee members, Professor Daniel Beard and Professor Peter Miller. Professor Beard has persuaded me to think deeply about mathematical research and to effectively communicate its interdisciplinary applications. Professor Miller inspired me to choose the field of nonlinear dynamics through his excellent course in my first semester, and has pushed me to strive for clarity and intention in my work through his curiosity and meticulousness.

I am grateful to the Department of Mathematics at the University of Michigan for providing a conducive environment for well-rounded personal and professional development, through multiple social events, research seminars and other growth opportunities. I would like to thank Professor Silas Alben for mentoring me throughout my Ph.D., acknowledging and appreciating my organizational work for the AIM student seminar and the SIAM minisymposiums. I would like to thank the very helpful administrative staff at the department, including Teresa Stokes, Kayla Papp and others, for all their efforts in making us feel welcomed and assisting our journey. I would like to thank our course coordinators, Elizabeth Wolf, Hanna Bennett and others, from whom I have learnt so much about teaching and communication.

I would like to extend my gratitude to Professor Warren Hare and Professor Ravindra B Bapat for instilling the initial curiosity and research discipline in me, also laying the foundation for my graduate work. I would also like to thank my managers at my internships, Mehrdad Yousefzadeh, Amal Chaturvedi, Kinjal Sanghavi and Eric Song, for their mentorship and for preparing me well for interdisciplinary industry.

I want to deeply thank my friends who have made this journey possible. I am truly indebted to all their support, love, compassion and joyful memories. To friends that I made in Michigan, Suchi, Urshita, Shivani, Paul, Christiana, Malavika, Shirlyn, Jiajia, Guanhua, Schinella, Han, Abhiti, Varun, Hrishi, Preetham, Tejaswi, Shravan, Cyril, Bernardo, Chris, Teresa, Katja, Oliver, and many more, thank you for resonating with me, laughing at my jokes, and creating a second home away from home. To friends I have back at home, Khyati and Samiksha, whose video calls and check-ins have kept me sane. To my best friends, Mehal and Sridhar, thank you for always believing in me even when I didn't, and showing up for me unconditionally. I am extremely lucky to have these pillars of support in my life.

Last and definitely not the least, I would like to thank my family, my parents and my sister for their support, sacrifices and utmost faith in my dreams. They pushed me to see the world and seek unimaginable pursuits and this wouldn't have been possible without them.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES viii

LIST OF TABLES xv

LIST OF APPENDICES xvi

LIST OF ACRONYMS xvii

ABSTRACT xix

CHAPTERS

1 Introduction 1
- 1.1 Mathematical Modeling for Enzyme Kinetics 3
  - 1.1.1 Singular Perturbations and Fenichel Theory 5
  - 1.1.2 Quantitative and Qualitative Analysis of Reduced Models 13
- 1.2 Mathematical Modeling for Sleep-Wake Regulation 15
  - 1.2.1 Neurobiology for Data Retrieval and Flip-Flop Models 16
  - 1.2.2 Methods and Measures for Parameter Estimation 20
- 1.3 Overview and Contributions 22

2 Stochastic Enzyme Kinetics and the Quasi-Steady-State Reductions: Application of the Slow Scale Linear Noise Approximation à la Fenichel 30
- 2.1 Introduction 30
- 2.2 Quasi-Steady-State Approximations of Michaelis–Menten Mechanism 31
  - 2.2.1 Standard Quasi-Steady-State Approximation 32
  - 2.2.2 Reverse Quasi-Steady-State Approximation 32
- 2.3 Stochastic Chemical Kinetics: Expansions, Reductions, and Approximations 35
  - 2.3.1 Stochastic Chemical Kinetics Far from the Thermodynamic Limit: The Master Equation 35
  - 2.3.2 Approaching the Thermodynamic Limit: the LNA 36
  - 2.3.3 Projecting onto the Tangent Space of the Critical Manifold 38
  - 2.3.4 The ssLNA for Systems in Standard Form: Comparison to Previous Results 40

2.3.5 Benchmark Example: The MM Reaction Mechanism in the Limit of Small k_2 in (c, p) Coordinates . . . 44
2.4 Reduction of the CME: Intimations from the Linear Noise Regime . . . 46
2.4.1 Dynamic Bifurcations and the Segel–Slemrod sQSS Condition . . . 46
2.5 Discussion . . . 53
3 Accuracy versus Predominance: Reassessing the Validity of the Quasi-Steady-State Approximation . . . 54
3.1 Introduction . . . 54
3.2 The Open Problem: Identifying the Predominant Reduction . . . 56
3.3 Anti-Funnels and the Slow Invariant Manifold . . . 60
3.4 Trapping the Slow Manifold: The Standard QSSA . . . 65
3.4.1 Accuracy of the Standard QSSA . . . 70
3.4.2 Validity of the sQSSA for High Enzyme Concentrations . . . 71
3.5 Trapping the Slow Manifold: The Slow Product QSSA . . . 74
3.5.1 Revisiting the Reich-Sel'kov Condition . . . 82
3.6 Discussion . . . 83
4 Gaining Hyperbolicity: A Blow-up Approach to Understanding Quasi-Steady-State Approximations . . . 85
4.1 Introduction . . . 85
4.2 Limitations of Fenichel theory . . . 88
4.2.1 Loss of Normal Hyperbolicity . . . 88
4.2.2 Quantitative Understanding of ε . . . 91
4.3 The Toolbox: Blow-up, Quasi-homogeneity and Coordinate Charts . . . 91
4.3.1 The Blow-up Map . . . 91
4.3.2 Quasi-Homogeneity of Vector Fields . . . 94
4.3.3 Defining Coordinate Charts using Directional Blow-up(s) . . . 96
4.4 The IAZA Reaction Mechanism . . . 97
4.4.1 Dynamics in (w, c) Phase-Plane . . . 97
4.4.2 Solution Trajectories in the Phase-Plane . . . 98
4.5 Loss of Normal Hyperbolicity: Dynamic Transcritical Bifurcation Point . . . 102
4.5.1 Reduced Systems in the Presence of Bifurcation . . . 103
4.5.2 Blow-up Analysis near the Bifurcation Point . . . 106
4.5.3 Validity of the Reverse QSS Approximation . . . 116
4.6 Dynamic Imperfection, Faux Canard and More . . . 120
4.6.1 The Irreversible IAZA Mechanism . . . 122
4.6.2 Validity of the Classical Reduction . . . 125
4.7 Discussion . . . 127
5 Gaining Space, not Hyperbolicity: A Blow-up Approach to Understanding Quasi-Steady-State Approximations . . . 129
5.1 Introduction . . . 129
5.2 Intermolecular Autocatalytic Zymogen Activation Kinetics for Small k_1 . . . 132
5.2.1 Phase-plane Dynamics and Quasi-Steady-State Reduction . . . 133
5.2.2 Blowing-up the Zymogen Axis . . . 134
5.3 Michaelis–Menten Kinetics for Small e_0 . . . 139
5.3.1 Phase-plane Dynamics and Quasi-Steady-State Reduction . . . 140

5.3.2 Blowing-up the Substrate Axis . . . 141
5.4 Discussion . . . 144
6 Gradient-based Optimization of Noisy ODE Models for Polyphasic Rodent Sleep-Wake Behavior . . . 147
6.1 Introduction . . . 147
6.2 Experimental Data: Sleep-Wake Behavior in Rodents . . . 149
6.2.1 Descriptive Measures of Sleep-Wake Data . . . 151
6.2.2 Underlying Distribution for Wake Activation . . . 152
6.3 Mathematical Model: Simulating the Sleep-Wake Dynamics . . . 154
6.3.1 Flip-Flop ODE Model . . . 154
6.3.2 Model Equations and Parameters . . . 156
6.3.3 Addition of Noise to ODE Model . . . 157
6.4 Methodology: Fitting Model to Data . . . 162
6.4.1 Gradient Descent using Automatic Differentiation . . . 163
6.4.2 Sensitivity Analysis for Model Parameters . . . 165
6.4.3 Data Processing and Augmentation . . . 168
6.4.4 Choice of Optimization Hyperparameters: Error, Kernel and Optimizer 170
6.4.5 Model Evaluation . . . 172
6.5 Results . . . 173
6.5.1 Model Dynamics are Driven by f_W Noise Stimuli and the Underlying Deterministic Behavior . . . 174
6.5.2 Refining Model Fit with Noisy $h_{\max}$ Inputs . . . 178
6.5.3 Fine-tuning Weakly Correlated Parameters Enhances Model Fit . . . 181
6.6 Discussion . . . 182
7 Conclusions and Future Work . . . 185

APPENDICES . . . 189

BIBLIOGRAPHY . . . 205

LIST OF FIGURES

FIGURE

1.1 **The slow flow for the MM mechanism converges to the standard QSS approximation as $e_0 \to 0$.** The dashed black curves are the numerical solutions to the full MM ODE model (1.2) for different initial conditions. The solid blue curve is the QSS curve for the standard Fenichel reduction (1.24). **Left**: $e_0 = 5$. **Right**: $e_0 = 0.1$. **Both**: The initial conditions are $s_0 \in \{2, 4, 6, 8, 10\}$ and $c_0 = 0$. The parameter values are $k_1 = 1, k_{-1} = 1, k_2 = 1$. 12

1.2 **The activity in presynaptic neuronal populations govern neurotransmitter concentrations which affect the activity in postsynaptic neuronal populations.** 18

1.3 **Time traces showing the variation of f_W, f_S and $h(t)$ over 2 hours**. The traces are obtained by numerically solving a deterministic ODE flip-flop system with initial conditions $f_W(0) = 3$, $f_S(0) = 0$, $h(0) = 0.24637$. The system is solved for every 1.25 seconds. All the details for the model are presented in Chapter 6. 19

2.1 **A dynamic transcritical bifurcation occurs in the singular limit corresponding to $k_2 = k_{-1} = 0$ of the MM reaction mechanism.** At $S^\star$ the Jacobian has a double-zero eigenvalue. Along the dashed lines the Jacobian has one zero eigenvalue and one positive eigenvalue. Along the solid lines the Jacobian has one zero eigenvalue and one negative eigenvalue. At the bifurcation point $S^\star = (0, e_0)$, the lines $S_a^r \cup S_r^l$ and $S_r^t \cup S_a^b$ intersect and exchange stability. 34

2.2 **The stochastic sQSSA (2.45) will fail near the bifurcation point if $e_0/K_M \gg 1$.** In both panels, the solid green curve is the numerically-computed mean of the timecourse for N (the number of product molecules) obtained from 1000 simulations generated by the Gillespie algorithm; the dashed/dotted green curves demarcate the mean $\pm$ the standard deviation. The solid black curve is the numerically-computed mean of the time course for N (the number of product molecules) obtained from 1000 simulations generated by the Gillespie algorithm equipped with a QSS-derived propensity function; the blue line shows one such randomly picked simulation; the dashed/dotted black lines demarcate the mean $\pm$ the standard deviation. In both simulations, the parameters (in arbitrary units) are: $n_{s_0} = 10000$, $n_{e_0} = 100$, $k_1 = 1000.0$, $k_{-1} = 0.01$, $k_2 = 0.01$ and $\Omega = 1.0$. $n_s(0) = n_{s_0}$, $n_c(0) = 0$, $n_e(0) = n_{e_0}$, and $N(0) = 0$. Note that $e_0/(K_M + s_0) \approx 0.01$ and therefore the Segel-Slemrod (2.47) condition holds. LEFT panel: The black solid and dashed/dotted lines are obtained from Gillespie algorithm equipped with the propensity function (2.57); the first two statistical moments are practically indistinguishable. RIGHT panel: The black solid and dashed/dotted lines are obtained from Gillespie algorithm equipped with the propensity function (2.45), and the stochastic sQSSA fails near $N = N^*$. . . . 52

3.1 **The sQSSA** (3.5) **and the slow product QSSA** (3.7) **are indistinguishable for large k_{-1}. Left**: $s_0 = 10$, $c_0 = 0$, $e_0 = 10$, $k_1 = 1$, $k_{-1} = 100$, $k_2 = 1$. **Right**: $s_0 = 10$, $c_0 = 0$, $e_0 = 10$, $k_1 = 1$, $k_{-1} = 500$, $k_2 = 1$. Substrate depletion over time is shown in both panels. $s_c(t)$ is the numerical solution to (3.5) and $s_{SP}(t)$ is the numerical solution to (3.7). 58

3.2 **The region, Γ, between the horizontal nullcline and the vertical nullcline curves is a trapping region on the phase plane.** The numerical solutions to the mass action equations (3.1) for several initial conditions are denoted by black dashed lines. $s_0 \in \{1, 2, 3, 4, 5, 6\}$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 1$, $k_2 = 1$. The numerical solution of trajectories start on the s-axis and eventually enter Γ. 62

3.3 **The region between the horizontal nullcline curve $\gamma_c(s)$ and the curve $\gamma(s)$ is a trapping region on the phase plane.** The numerical solutions to the mass action equations (3.1) for several initial conditions are denoted by black dashed lines. $s_0 \in \{1, 2, 3, 4, 5, 6\}$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 1$, $k_2 = 1$. 67

3.4 **The slow product QSS curve γ_{SP} lies above the horizontal nullcline γ_c for $s < s^*$** (3.50) **in the phase plane.** The blue asterik denotes the intersection point at s^* (3.50). **Top Left ($k_2 > e_0 k_1$)**: $k_1 = 0.1$, $k_2 = 1$, $k_{-1} = 1$, $e_0 = 6$, $s^* = -4$. **Top Right ($k_2 = e_0 k_1$)**: $k_1 = 0.1$, $k_2 = 0.6$, $k_{-1} = 1$, $e_0 = 6$, $s^* = 0$. **Bottom Left ($k_2 < e_0 k_1$)**: $k_1 = 1$, $k_2 = 1$, $k_{-1} = 1$, $e_0 = 6$, $s^* = 5$. **Bottom Right ($k_2 < e_0 k_1$)**: $k_1 = 1$, $k_2 = 0.4$, $k_{-1} = 1$, $e_0 = 6$, $s^* = 14$. 75

3.5 **The region between the slow product QSS curve $\gamma_{SP}(s)$ and the curve $\gamma(s)$ is a trapping region for solutions in the phase plane when $e_0 < 8K_S$.** The intersection point of γ_{SP} and γ_c, s^*, is denoted by the blue asterik. The numerical solutions to the mass action equations (3.1) for several initial conditions are denoted by black dashed lines. $s_0 \in \{2, 6, 12\}$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 1$, $k_2 = 1$. 76

3.6 **The condition $e_0 < K$ ensures that the sQSSA provides a better approximation than the slow product QSSA under the Reich-Sel'kov condition.** $e_0 \ll K_M$ is satisfied in both panels. **Left ($e_0 < K$)**: $s_0 = 9$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 100$, $k_2 = 10$. **Right ($e_0 > K$)**: $s_0 = 9$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 100$, $k_2 = 4$. Substrate depletion over time is shown in both panels. The numerical solution to the mass action equations (3.1) is denoted by black dashed lines. $s_c(t)$ is the numerical solution to (3.5) and $s_{SP}(t)$ is the numerical solution to (3.7). 82

4.1 **The movement of solution trajectories in the phase-plane when $z_0 > e_0$ and $z_0 < e_0$.** The solid black curve is the numerical solution to the full IAZA ODE model (4.7) and the dashed red curve is the numerical solution to the reverse-QSSA (4.44) in the (w, c) phase-plane. **Left**: The parameter values are $k_1 = 10, k_{-1} = 1, k_2 = 1, z_0 = 10, e_0 = 5$. **Right**: The parameter values are $k_1 = 10, k_{-1} = 1, k_2 = 1, z_0 = 5, e_0 = 10$. 89

4.2 The phase plane geometry of the IAZA mechanism in (w, c) coordinates contains three nullclines. The w-nullcline is given by the w-axis depicted in red. The c-nullcline is given by (4.21) depicted by the two blue curves. The curve that intersects with the w-axis is $c = h^-(w)$. The two equilibrium points are at $(-e_0, 0)$ and $(z_0, 0)$ which are saddle and stable points respectively. **Left**: $e_0 \leq z_0$. The complex concentration of a solution trajectory is bounded above by $h^-((z_0 - e_0)/2)$. **Right**: $e_0 > z_0$. The complex concentration of a solution trajectory is bounded above by $h^-(0)$. 98

4.3 **The (w, c) phase-plane inhabitates a dynamic transcritical bifurcation point at the intersection of Γ_1 and Γ_2 (4.9) given by (w^*, c^*) (4.41).** The solid red lines denote the stable submanifolds and the dashed red lines denote unstable submanifolds. 103

4.4 **The blown-up space can be divided into 3 coordinate charts.** The blown-up space is defined for coordinates $(\bar{w}, \bar{c}, \bar{\varepsilon}, r)$ and corresponds to the space $S^2 \times [0, r_0]$. Chart K_1 covers the region where trajectories enter. Chart K_2 covers a small neighborhood of the bifurcation point at the origin. Chart K_3 covers the region where trajectories exit. The origin in (w, c, ε) is mapped to $\bar{B}_0 := S^2 \times \{0\}$ for $r = 0$. 107

4.5 **The zeroth-order solution estimate for the desingularized blown-up system improves as r_2 (or equivalently ε) decreases.** The solid black line is the numerical solution to system (4.49) with singularity $r_2 = \sqrt{\varepsilon}$. The dashed black line is the solution to the desingularized zeroth-order system (4.69). The red dashed line is the numerical solution to the reverse QSSA blown-up to the chart K_2. The green dashed line is the numerical solution to the classical QSSA blown-up to the chart K_2. **Both**: Parameter values used are $z_0 = 10, e_0 = 5, k_1 = 1, k_{-1} = 1, k_2 = 1$, Initial Conditions used are $w(0) = -(z_0 - e_0)/(2r_2), c(0) = -E_T/(2r_2)$. **Left**: $r_2 = 0.1$. **Right**: $r_2 = 0.05$. 116

4.6 **The peak point estimate** (4.76) **for the solution trajectory improves as ε decreases.** The solid black line is the numerical solution to system (4.39). The red dashed line is the numerical solution to the reverse QSSA (4.44). The green dashed line is the numerical solution to the classical QSSA blown-up (4.45). **Both**: Parameter values used are $z_0 = 10, e_0 = 5, k_1 = 1, k_{-1} = 1, k_2 = 1$, Initial Conditions used are $w(0) = 0, c(0) = 0$. **Left**: $\varepsilon = 0.01$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (1.29 \times 10^{-3}, 1.51 \times 10^{-3})$; **Right**: $\varepsilon = 0.0025$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (3.07 \times 10^{-4}, 3.71 \times 10^{-4})$. 119

4.7 **The peak point estimate** (4.86) **for the solution trajectory for irreversible IAZA mechanism improves as ε decreases.** The solid black line is the numerical solution to system (4.84). The red dashed line is the numerical solution to the reverse QSSA (4.44). The green dashed line is the numerical solution to the classical QSSA blown-up (4.45). **Both**: Parameter values used are $z_0 = 10, e_0 = 5, k_1 = 1, k_{-1} = 0, k_2 = 1$, Initial Conditions used are $w(0) = 0, c(0) = 0$. **Left**: $\varepsilon = 0.01$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (2.06 \times 10^{-3}, 8.66 \times 10^{-4})$; **Right**: $\varepsilon = 0.0025$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (6.72 \times 10^{-5}, 2.14 \times 10^{-4})$. 123

4.8 **The irreversible IAZA mechanism conforms with the normal form of a transcritical bifurcation and admits faux canard in the phase-plane.** For the same initial conditions, the reversible IAZA mechanism does not exhibit canard-like behavior. The solid black line is the numerical solution to system (4.39). The red dashed line is the numerical solution to the reverse QSSA (4.44). **Both:** The parameter values are $z_0 = 10, e_0 = 5, k_1 = 1, k_2 = 1, \varepsilon = 0.01$, The initial conditions are $w(0) = 0.5, c(0) = 9.5$. **Left:** The irreversible case corresponds to $k_{-1} = 0$. **Right:** The reversible case corresponds to $k_{-1} = 1$. . . 124

4.9 **As k_{-1} increases, the classical nullcline-based reduction provides more accurate approximation as compared to the reverse QSSA for the IAZA mechanism.** The solid black line is the numerical solution to system (4.39). The red dashed line is the numerical solution to the reverse QSSA (4.44). The green dashed line is the numerical solution to the classical QSSA blown-up (4.45). **Both:** The parameter values are $z_0 = 10, e_0 = 5, k_1 = 1, k_2 = 1, \varepsilon = 0.01$, The initial conditions are $w(0) = 0, c(0) = 0$. **Left:** $k_{-1} = 1$. **Right:** $k_{-1} = 10$. Both panels are plotted on the same axis range and scale to highlight the difference in approximation accuracy. 126

5.1 **The IAZA mechanism behaves remarkably like a system in standard form when k_1 is small.** The dashed black curves are the numerical solutions to the full IAZA ODE model (5.7) for different initial conditions. The solid blue curve is the QSS curve for the standard Fenichel reduction (5.10) for a fixed $E_T = 15$. **Left**: The parameter values are $k_1 = 0.1, k_{-1} = 1, k_2 = 1$. **Right**: The parameter values are $k_1 = 0.001, k_{-1} = 1, k_2 = 1$. **Both**: The initial conditions are $z_0 \in \{2, 4, 6, 8, 10\}$, $c_0 \in \{0, e_0\}$ and $e_0 = E_T - z_0$ such that $E_T = 15$. 133

5.2 **The blown-up space can be divided into 2 coordinate charts.** The blown-up space is defined for coordinates $(z, \bar{c}, \bar{\varepsilon}, r)$ and the cylinder corresponds to the space $\mathrm{S}^1 \times \mathbb{R} \times [0, r_0]$. Chart K_1 is an intermediary blow-up in the direction of $\bar{\varepsilon}$. Chart K_2 is an intermediary blow-up in the direction of $\bar{c}$. The origin in (w, c, ε) is mapped to $\bar{Q}_0 := \mathrm{S}^1 \times \mathbb{R} \times \{r = 0\}$. 135

5.3 **The MM mechanism behaves remarkably like a system in standard form when e_0 is small.** The dashed black curves are the numerical solutions to the full MM ODE model (5.27) for different initial conditions. The solid blue curve is the QSS curve for the standard Fenichel reduction (5.31). **Left**: $e_0 = 5$. **Right**: $e_0 = 0.1$. **Both**: The initial conditions are $s_0 \in \{2, 4, 6, 8, 19\}$ and $c_0 \in \{0, e_0\}$. The parameter values are $k_1 = 1, k_{-1} = 1, k_2 = 1$. 141

6.1 **A hypnogram depicting 1 hour of wake, REM and NREM states from sleep scored data.** Unpublished data is courtesy of the Weber Lab (U of Pennsylvania); The mouse ID is $c50_042418n1$; and the hour shown is from 7 AM to 8 AM. Mice in this dataset are exposed to a 12/12 light/dark schedule with lights turned on at 7 AM. 149

6.2 **A hypnogram depicting 1 hour of sleep/wake states by combining NREM and REM together from sleep score data.** We also relabel wake recordings with 1 and sleep recordings as 0. The mouse ID is $c50_042418n1$ and the hour shown is from 7 AM to 8 AM. 150

6.3 **The six inference measures vary significantly over the two observation phases.** We have removed the outlier recordings and the 6 dark phase recordings which contain only sleep. We see a negative range of mean wake durations for the dark phase possibly because there is a huge variation in the mean wake measure for the dark phase recordings. 151

6.4 **The probability density function and the cumulative density function of the best fit theoretical exponential distribution overlap with the empirical data of sleep durations.** We use the python package 'distfit' for these plots. The x-axis denotes the values of sleep durations for the 50 light-phase data recordings from 9 AM to 11 AM. The best fit exponential distribution has the parameters loc = 0 and scale = 78.549. This corresponds to a probability density function of $\mathrm{pdf}(x) = \exp(-(x - \mathrm{loc}))/\mathrm{scale}$. In terms of the rate parameter, r, this corresponds to $r = 1/\mathrm{scale}$. 153

6.5 **A schematic representation for the physiological two-state model.** The wake and sleep promoting neuronal populations correspond to the neurotransmitters Noradrenaline (NE) and Gamma Aminobutyric Acid (GABA) respectively. The circles denote inhibitory synaptic connections between the two populations. h denotes the homeostatic input and Noise denotes the noise stimuli input to the wake-promoting population. 154
6.6 **Time traces showing the variation of f_W, f_S, $h(t)$ and Noisy stimuli over 2 hours**. The traces are obtained by numerically solving the ODE system with modified (6.6) and initial conditions $f_W(0) = 3$, $f_S(0) = 0$, $h(0) = 0.24637$. The system is solved for every 1.25 seconds. The value of r is fixed at 0.003. . . 159
6.7 **Time traces showing the variation of f_W, f_S, $h(t)$ and Noisy h_{max} over 2 hours.** The traces are obtained by numerically solving the ODE system with modified (6.7) and initial conditions $f_W(0) = 3$, $f_S(0) = 0$, $h(0) = 0.24637$. The system is solved for every 1.25 seconds. The value of r_h is fixed at 0.003, amplitude for $h_{max} = 20$ and mean for $h_{max} = 1$. 161
6.8 **Pairwise plots between k_2 and other model parameters show correlated wake duration variation in the deterministic model.** We consider a $\pm 20\%$ perturbation from the baseline values of the parameters. 166
6.9 **Pairwise plots between k_1 and other model parameters show correlated sleep duration variation in the deterministic model.** We consider a $\pm 20\%$ perturbation from the baseline values of the parameters. 167
6.10 **The six inference measures show similarity over the two hour time blocks.** Block 1 corresponds to 50 mice recordings from 9 AM to 11 AM. Block 2 corresponds to 50 mice recordings from 11 AM to 1 PM. See C.1 for sample-level details. 169
6.11 **Survival curves overlap for wake and sleep durations collected for light phase data over 9 AM-11AM and 11AM-1PM.** This motivates the augmentation of the two sets to collate a light phase dataset of 100 mice recordings. 169
6.12 **Inference Measures show some discrepancies for optimized model with noisy stimuli input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameters. 174
6.13 **Population-level survival curves and histograms show some discrepancies for optimized model with noisy stimuli input.** While the shape of the sleep distribution is captured by the model output, the wake distribution lacks the elongated shape admitted by the data. Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameters. . . . 176

6.14 **Inference Measures align better and admit more sample-level variability for optimized model with noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12), other baseline parameters with appended noisy h_{max}. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13), other baseline parameters with appended noisy h_{max}. . . 178
6.15 **Population-level survival curves and histograms admit similar shapes for optimized model with noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12), other baseline parameters with appended noisy h_{max}. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13), other baseline parameters with appended noisy h_{max}. 179
6.16 **Sample-level similarities between data and model improve for optimized model with both noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized Parameters denote the model with (6.12), (6.13), (6.14) and noisy stimuli. Optimized Parameters & Noisy h_{max} denotes the fully optimized noisy model with appended noisy h_{max}. 180
6.17 **Population-level similarities between data and model improve for optimized model with both noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized Parameters denote the model with (6.12), (6.13), (6.14) and noisy stimuli. Optimized Parameters & Noisy h_{max} denotes the fully optimized noisy model with appended noisy h_{max}. 181

C.1 **Exponential distribution provides good fit for the distribution of sleep durations in light phase data.** Other distributions such as beta, gamma and pareto also provide good fit. We use the python package 'distfit' for the plot. . . 204
C.2 **The Quantile-Quantile (QQ) plot compares the empirical distribution of sleep durations with the best fit theoretical exponential distribution.** It shows a good fit for the shorter durations of sleep but deviates from the 45-degree line. We use the python package 'distfit' for the plot. 204

LIST OF TABLES

TABLE

6.1 **Baseline Parameter Values for two-state flip-flop ODE model with exogenous noise.** α_P and β_P are in units of effective synaptic input. $g_{n,P}$ has units of effective synaptic input/Hz. h_{max} and h_{min} have units of percentage mean slow wave activity (SWA). k_1 and k_2 have units of effective synaptic input and effective synaptic input/(% mean SWA), respectively. 158

6.2 **Maximum Mean Discrepancy improves as we utilize noisy h_{max} for the optimized models.** The values denote the MMD measure computed between 100 data samples and 100 model outputs. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameters. Optimized Weak Parameters refers to model with (6.12), (6.13) and (6.14) and other baseline parameters. 173

6.3 **Wasserstein distance improves as we utilize noisy h_{max} for our converged models.** The values denote the Wasserstein distance computed between the collective wake and sleep durations observed for 100 data samples versus 100 model outputs. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameter. Optimized Weak Parameters refers to model with (6.12), (6.13) and (6.14) and other baseline parameters. 177

LIST OF APPENDICES

A Appendix for Chapter 3 189
A.1 Singular perturbation analysis of large k_{-1} 189
A.2 QSS approximations and initial conditions 190
B Appendix for Chapter 4 194
B.1 Quasi-Homogeneity 194
B.2 Coordinate Charts for a Blow-Up Analysis 194
C Appendix for Chapter 6 197
C.1 Inference Measures for Experimental Data 197
C.2 Theoretical Distributions for the Sleep Data 203

LIST OF ACRONYMS

BF basal forebrain

CME chemical master equation

DR dorsal raphe

ECG electrocardiogram

EEG electroencephalogram

EMG electromyography

GABA gamma-amino butyric acid

GSPT geometric singular perturbation theory

IAZA Intermolecular Autocatalytic Zymogen Activation

LC locus coeruleus

LDT laterodorsal tegmental nuclei

LSE least-squares estimation

MAP maximum a posteriori estimation

MCMC Markov Chain Monte Carlo

MLE maximum likelihood estimation

MM Michaelis–Menten

MMD maximum mean discrepancy

MnPO median preoptic nuclei

NE noradrenaline

NN neural network

NREM Non-Rapid Eye Movement

ODE ordinary differential equation

PDE partial differential equation

PPT pedunculopontine

QSSA Quasi-steady-state approximation

REM Rapid Eye Movement

RMSE root mean square error

rQSSA reverse quasi-steady-state approximation

SCN suprachiasmatic nucleus

SDE stochastic differential equation

sQSSA standard quasi-steady-state approximation

ssLNA slow scale linear noise approximation

SWA slow wave activity

TFPV Tikhonov-Fenichel parameter value

TMN tuberomammillary nucleus

VLPO ventrolateral preoptic nucleus

WENDy Weak form Estimation of Nonlinear Dynamics

ABSTRACT

Reduced models for enzymatic mechanisms, widely used in real-world biochemical applications, are often governed by parametric restrictions that determine their accuracy and applicability. The work in this thesis focuses on addressing challenges in analyzing solutions of reduced enzyme kinetic models, particularly in settings where assumptions of traditional analytic methods may fail. A second focus of this thesis is the development of a gradient-based optimization framework for fitting noisy differential equation models for the neuronal regulation of sleep-wake behavior to biological data.

In Chapter 2, we derive slow-scale linear noise approximation (ssLNA) to stochastic enzyme kinetic models directly using geometric singular perturbation theory (GSPT). We resolve discrepancies among prior stochastic reductions by identifying the role of the Tikhonov standard form in producing consistent reductions. Extending the ssLNA to systems with a transcritical bifurcation, we disprove the widely-used Segel–Slemrod condition for stochastic model reduction in the Michaelis–Menten (MM) mechanism. Instead, we demonstrate that the more restrictive Reich-Sel'kov condition is needed, aligning stochastic and deterministic validity thresholds for quasi-steady-state approximations (QSSAs).

In Chapter 3, we focus on the MM model to investigate the parametric conditions under which the standard QSSA is not only valid but also predominant among competing reductions. Using the theory of fences and anti-funnels, we locate the slow manifold in the MM phase plane and derive sharp error bounds. While the Reich-Sel'kov condition governs accuracy, we introduce a new and more restrictive qualifier that predicts the predominance of the standard QSSA. This distinction is critical when multiple reduced models are valid, as it guides optimal model selection for experimental design.

In Chapter 4, we perform model reduction analysis for the Intermolecular Autocatalytic Zymogen Activation (IAZA) model, when it features a dynamic transcritical bifurcation. GSPT fails in this setting due to loss of normal hyperbolicity. Using the blow-up method, we transform the system near the bifurcation to recover hyperbolicity and derive an explicit solution to the desingularized system at the bifurcation point. This allows us to estimate the peak complex concentration analytically and explain the surprising accuracy of the nullcline-based reduction. We also identify the existence of a canard and analyze the implications of dynamic imperfections for model reduction.

Chapter 5 tackles perturbation systems that lack an explicit Tikhonov standard form structure but mimic typical standard form behavior portraying fast convergence to QSS approximations. We propose a novel use of the blow-up method to connect the fast and slow dynamics in systems where a clear distinction is not apparent. Applied to the MM and IAZA systems, this approach reveals that the systems admit latent standard fast-slow structure and the trajectories starting both on and away from the critical manifold get attracted to the QSS curves, validating the resemblance to standard form dynamics.

Chapter 6 shifts focus to computational neuroscience and presents a framework for fitting noisy differential equation models to rodent sleep-wake data using gradient-based optimization. We adapt a mechanistic model of interacting neuronal populations with noisy firing activity. Leveraging automatic differentiation, maximum mean discrepancy, and sensitivity analysis, we demonstrate efficient parameter inference from discrete sleep score data. The methodology offers a framework to relate experimental hypotheses for the neuronal circuit structures responsible for sleep-wake regulation to the random polyphasic sleep-wake behavior of rodents, with applications in and beyond sleep modeling.

CHAPTER 1

Introduction

Mathematical models of biological processes provide valuable insights and propel theoretical understanding and practical applications of numerous cellular and physiological systems. For example, the study of enzymatic chemical reactions that occur in cells and tissues informs decisions in pharmaceutical research and drug discovery [77]. Numerous significant factors such as drug dosage and timing, patient-level variability and survival probabilities are all investigated using mathematical models that describe the associated chemical processes [32, 78]. In neuroscience, models that explicate neurological pathways have contributed to the understanding of brain processes and cognition and enabled advances in understanding and treating neuro-pathologies such as epilepsy, Parkinson's disease, and sleep-related disorders [30, 83, 185, 128]. These diverse applications underscore the need for accurate, interpretable, and computationally tractable models of chemical and biological mechanisms.

Mathematical models for enzyme kinetics consist of systems of ordinary differential equations (ODEs) governing the concentrations of chemical species over time. Due to their complexity, non-linear nature and high dimensionality, full-scale models of chemical systems are often intractable for analysis. Reduced models of chemical mechanisms are therefore used in experimental design, aiding real-world applications such as pharmaceutical research. Deepening our understanding of the validity and accuracy of these reduced models becomes crucial. Popular reduction theories, such as geometric singular perturbation theory, provide theoretical justification for model simplification [54, 184]. Moreover, specific biochemical parametric conditions have been identified that guarantee the reliability of the reduced mod-

els [160, 49]. However, several open questions remain. Can reduced models also offer insight into the stochastic regime which resembles real-world noisy scenarios? What happens when multiple reduced models are simultaneously valid and yield comparably accurate results? Is there a principled way to select the most accurate reduced models for optimal performance? Additionally, how do we proceed when the assumptions required for the classical reduction theories break down? In this thesis, we address these questions for two important chemical mechanisms and propose generalizable frameworks. This work contributes to strengthening the theoretical foundation and practical reliability of model reduction in enzyme kinetics and beyond.

Another focus of this thesis is on modeling the neurophysiological systems that regulate transitions between sleep and wake states. Sleep is a fundamental physiological process that influences cognitive performance, immune function, and overall well-being [58]. Computational neuroscience provides a powerful framework to study the neural circuits and biochemical pathways governing these sleep-wake transitions. Through mechanistic differential equations based models grounded in neurobiology, we can capture the dynamical interactions between wake-promoting and sleep-promoting neuronal populations [16]. These models not only help to clarify how sleep-wake cycles emerge from circuit-level activity, but also offer predictive tools for analyzing experimental data, optimizing interventions, and generating testable hypotheses. In particular, simplified yet biophysically informed models, such as those employing firing-rate model formalisms, enable tractable analyses while retaining essential features of the system [34, 39]. The work in this thesis is motivated by the need for accurate models that bridge the gap between theory and experimental sleep data, particularly in the context of estimating physiological parameters.

In the following sections, we provide a brief literature review for the two topics of focus: enzyme kinetics and sleep-wake regulation. This includes an overview of the popular Fenichel reduction theory used to obtain reduced models of chemical mechanisms along with conditions for their validity in experimental applications. We then review the neurobiology

of sleep-wake regulation and associated ordinary differential equation (ODE) models based on current hypotheses of the underlying brain dynamics. We also detail some past works on parameter estimation in models of biological systems. We conclude this chapter with an overview of the thesis and our main contributions.

1.1 Mathematical Modeling for Enzyme Kinetics

Enzyme kinetics refers to the study of the rates of chemical reactions that involve enzymes. An enzyme is a molecule that acts as a biological catalyst to accelerate a chemical reaction. The simplest case of an enzyme-catalysed chemical reaction is given by

$$\mathrm{S} + \mathrm{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_2} \mathrm{E} + \mathrm{P}, \tag{1.1}$$

referred to as the Michaelis–Menten system, a cornerstone of enzyme kinetics [73, 118, 20]. Here, the Enzyme binds to the Substrate to form an Enzyme-Substrate intermediary complex, C. The formation of the Complex is a reversible process. Finally, the Complex transforms to the Product molecule and recovers the Enzyme molecule. The constants, k_1, k_{-1} and k_2 are the reaction rates. Mechanism (1.1), initially modeled to describe the hydrolysis of sucrose, explains various enzyme-substrate binding processes ranging from protein-protein interaction to antigen-antibody binding and more [27, 139].

Enzymatic chemical mechanisms have been studied through a range of deterministic, numerical and stochastic models. When the chemical species are present in abundance i.e., in the thermodynamic limit, these processes are typically modeled using nonlinear ordinary differential equations [52]. This approach is referred to as the *deterministic* modeling of chemical mechanisms.

Typically, a model for a single chemical mechanism is comprised of various elementary reactions with multiple chemical species and parameters involved. For instance, the mass-

action equations describing the Michaelis–Menten mechanism (1.1) include,

$$\begin{aligned}
\dot{s} &= -k_1 es + k_{-1} c, \\
\dot{c} &= \quad k_1 es - (k_{-1} + k_2) c, \\
\dot{e} &= -k_1 es + (k_{-1} + k_2) c, \\
\dot{p} &= \quad k_2 c,
\end{aligned}$$

where s, c, e and p denote the concentrations of S, C, E and P, respectively and "$\dot{x}$" denotes the differentiation of x with respect to time t. Biologically relevant initial conditions are $s(0) = s_0$, $c(0) = 0$, $e(0) = e_0$ and $p(0) = 0$. Note that $\dot{c} + \dot{e} = 0$ and the quantity $c + e$ is conserved at e_0. Once the dynamics of $s(t)$ and $c(t)$ are known, the evolution of product $p(t)$ is easily recovered. Even then, the conservation laws lead to a two-dimensional system of nonlinear ordinary differential equations,

$$\dot{s} = -k_1 (e_0 - c) s + k_{-1} c, \tag{1.2a}$$

$$\dot{c} = \quad k_1 (e_0 - c) s - (k_{-1} + k_2) c. \tag{1.2b}$$

Closed-form solutions to this system remain unattainable due to quadratic nonlinearities. Similarly, the mathematical models that describe other simple chemical mechanisms soon become unwieldy and more importantly, hard to apply in real-world applications. This is where model reduction comes into play.

Continuing the same example, the standard quasi-steady-state approximation (sQSSA) to the mass action equations associated with the MM reaction mechanism is a well-known and extensively studied reduced model in biochemical kinetics and pharmaceutical applications [158, 160, 28]. Simply put, the sQSSA is a reduced system that approximates the depletion

of substrate over time as follows:

$$\dot{s} = -\frac{k_2 e_0 s}{K_M + s}, \quad s(0) = s_0, \tag{1.3}$$

where $K_M = (k_{-1} + k_2)/k_1$ is the Michaelis constant and $k_2\, e_0$ is the limiting rate. The Michaelis constant is composite of the equilibrium constant $K_S = k_{-1}/k_1$ and the Van Slyke-Cullen constant $K = k_2/k_1$, and can be expressed as $K_M = K_S + K$. The closed-form solution to the approximation (1.3) is known, rendering it highly advantageous for theoretical and experimental research [159]. Importantly, while a reduced model is an approximation to the full scale model, it is only valid for certain restricted biochemical and parametric conditions. Thus, we adhere to Fenichel theory that mathematically justifies the approximation and convergence of reduced models, to a certain extent.

In the following subsections, we provide the literature background for the Fenichel reduction theory, its assumptions and its implications. We also present detailed steps to derive the reduced model (1.3) from the full model (1.2) for a deeper mathematical understanding. Finally, we discuss previous works on the validity and accuracy of reduced models, specifically pertaining to chemical mechanisms and Fenichel theory.

1.1.1 Singular Perturbations and Fenichel Theory

A biochemical mechanism composed of elementary reactions that occur on disparate timescales is an example of a multiscale process. One of the most widely-studied forms of multiscale processes include two timescales, a slow and a fast timescale. The presence of widely separated timescales provides a way to systematically reduce the number of model equations that simulate the dynamics over the slow (long) timescale. Naturally, due to the reduced number of equations and parameters involved, reduced models are highly advantageous in experimental study of chemical mechanisms and in solving inverse problems like kinetic parameter estimation [24, 158]. The reduced systems approximate the dynamics of the original systems over a restricted range of parameter values and thus, all experimental

conditions for their applications must also align with their validity constraints [168]. As a result, a crucial component of the field of model reduction is the study of the validity of these reductions.

To this end, extensive research has been done on a reduction theory that identifies reduced systems and justifies their applicability. Fenichel theory, also known as Geometric Singular Perturbation Theory (GSPT), is one of the most widely used reduction theories for analyzing singularly perturbed dynamical systems. It provides a framework to obtain reductions known as *quasi-steady-state* reductions under specific biochemical limits [54, 55]. In this section, we present a brief overview of singularly perturbed ODE systems, Fenichel theory and some benchmark examples of the theory's applications.

Definition 1.1. Consider a perturbed dynamical system

$$\dot{x} = f(x, \varepsilon), \quad f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \tag{1.4}$$

where $\varepsilon \in \mathbb{R}$ is a parameter that is "small" in the sense that it is close to zero[1]. If the set of equilibrium points for $\varepsilon = 0$,

$$S_0 := \{x \in \mathbb{R}^n : f(x, 0) = 0\}, \tag{1.5}$$

is a non-empty k dimensional submanifold of $\mathbb{R}^n$ where $0 < k < n$, then (1.4) is said to be *singularly perturbed*.

Remark 1.1. The system (1.4) represents a general form of singularly perturbed systems. If the system is written with an explicit split between the fast reactions versus the slow

[1]We assume $f(x, \varepsilon)$ is sufficiently smooth in both arguments.

reactions such as

$$\dot{z} = f(z, y, \varepsilon), \qquad z \in \mathbb{R}^k,\ f \in \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^k, \tag{1.6a}$$

$$\dot{y} = \varepsilon g(z, y, \varepsilon), \qquad y \in \mathbb{R}^{n-k},\ g \in \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n-k}, \tag{1.6b}$$

it is known as a system in the standard fast-slow form defined by the popular Tikhonov's theorem [173]. As expected, most systems are present in the general form but the Fenichel theory aligns well with systems present in standard form.

In the case of a singularly perturbed system, the rank of the Jacobian with respect to x, $D_1 f(x, 0)$, will be constant and equal to $n - k$ along S_0. An important result from Fenichel theory pertains to the presence of normal hyperbolicity:

Definition 1.2. If the $n - k$ non-zero eigenvalues of $D_1 f(x, 0)$ are strictly bounded away from the imaginary axis everywhere along S_0, then S_0 is normally hyperbolic.

Fenichel theory states that if S_0 is compact and normally hyperbolic, then S_0 "persists", meaning that there exists an invariant manifold, S_0^ε, for $\varepsilon \in [0, \varepsilon_0)$. In the fortunate event that we have an explicit representation of S_0^ε as a graph[2] over the first k coordinates of x and ε,

$$S_0^\varepsilon := \{(z, y, \varepsilon) \in \mathbb{R}^k \times \mathbb{R}^{n-k} \times \mathbb{R} : y = h(z, \varepsilon)\}, \tag{1.7}$$

where $x := (z, y)$ and $z \in \mathbb{R}^k,\ y \in \mathbb{R}^{n-k}$, then the flow on S_0^ε is given by

$$\dot{z}_i = f_i(z, h(z, \varepsilon), \varepsilon), \quad 1 \le i \le k \tag{1.8}$$

where the f_i's denote the component functions of f. Since S_0^ε is a k dimensional manifold, it is homeomorphic to $\mathbb{R}^k$ in a neighborhood of every point belonging to S_0^ε; hence, only k coordinates are required to parameterize S_0^ε. Further, note that (1.8) describes the temporal evolution of k independent variables. The evolution of the remaining $n-k$ variables depends

[2]Such a representation is always locally possible.

completely on the evolution of z:

$$\dot{y} = D_1 h(z, \varepsilon)\dot{z}, \quad D_1 h(z, \varepsilon) \in \mathbb{R}^{(n-k)\times k}. \tag{1.9}$$

In singular perturbation theory, the submanifold S_0 comprised of stationary points is known as the *critical manifold*. The persistence of a normally hyperbolic S_0 with respect to a perturbation not only asserts the existence of a normally hyperbolic invariant manifold, S_0^ε, but also the persistence of S_0's stable and unstable manifolds, $W^s(S_0)$ and $W^u(S_0)$ concerning the same perturbation. This latter point is important because it says that S_0^ε inherits the same stability properties as S_0 (i.e., depending on S_0, S_0^ε will either be attractive $W^u(S_0) = \emptyset \implies W^u(S_0^\varepsilon) = \emptyset$, repulsive $W^s(S_0) = \emptyset \implies W^s(S_0^\varepsilon) = \emptyset$, or will be of saddle type and come equipped with both stable and unstable manifolds). If S_0 is both normally hyperbolic and attracting – meaning that the real parts of the $n - k$ nontrivial eigenvalues of $D_1 f(x, 0)$ are strictly negative, then (1.8) describes the long-time evolution of (1.4), since trajectories starting near S_0^ε will contract towards S_0^ε at an exponential rate. Naturally, the loss of normal hyperbolicity for the critical manifold poses problems for the reduction theory. We address this in chapter 4.

From the perspective of applications, the caveat to Fenichel's earlier results concerning the persistence of normally hyperbolic critical manifolds is that S_0^ε is usually unknown; we only know S_0 [54]. The question then becomes: Can we *build* an approximation to (1.8) without explicit knowledge of S_0^ε? The answer to this question was later addressed by Fenichel [55] himself. Expanding $f(x, \varepsilon)$ in a Taylor series near $\varepsilon = 0$ yields

$$\dot{x} = f(x, 0) + \left.\frac{\partial f(x, \varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} \cdot \varepsilon + \mathcal{O}(\varepsilon^2) =: f_0(x) + \varepsilon f_1(x) + \mathcal{O}(\varepsilon^2). \tag{1.10}$$

If S_0 is normally hyperbolic[3], then there exists a splitting

$$T_x\mathbb{R}^n \cong \mathbb{R}^n = \ker D_1 f(x,0) \oplus \text{ image } D_1 f(x,0) \tag{1.11}$$

for all $x \in S_0$. Note that $T_x S_0 \cong \ker D_1 f(x,0)$ for all $x \in S_0$. To approximate the flow on $\mathcal{S}_0^\varepsilon$, project the right hand side of (1.10) onto the tangent bundle of S_0, $T_x S_0$. Due to the decomposition, there exists a (unique) projection operator $\Pi^{S_0} : \mathbb{R}^n \to \ker D_1 f(x,0)$, which projects vectors in $\mathbb{R}^n$ onto $\ker D_1 f(x,0)$ along the image of $D_1 f(x,0)$ (i.e., $\text{id.} - \Pi^{S_0} : \mathbb{R}^n \to$ image $D_1 f(x,0)$, where "id." denotes the $n \times n$ identity matrix). The projection of (1.10) onto $T_x S_0$ is then given by the n-dimensional dynamical system (with k independent variables)

$$x' = \Pi^{S_0} f_1(x)|_{x \in S_0} \tag{1.12}$$

where "x'" denotes differentiation of x with respect to the slow timescale, $\tau = \varepsilon t$. More importantly, the flow on S_0^ε given by (1.8) converges to (1.12) as $\varepsilon \to 0$. The reduction (1.12) is referred to as quasi-steady-state approximation (QSSA) in chemical kinetics.

Generally speaking, solutions to (1.4) with an initial condition $x(0) = (z,y)(0) = (z_0, y_0)$ will not converge (as $\varepsilon \to 0$) to the solution generated from (1.12) equipped with the initial condition $z(0) = z_0$ unless z_0 happens to lie sufficiently close to S_0 or z and y are special and respectively constitute "slow" and "fast" coordinates. We address this caveat in chapter 5.

The traditional methodology utilized for the task of putting the mass action equations into perturbation form involves scaling and dimensional analysis. Such analyses are usually a mixture of logic and guesswork and, unless great care is taken, scaling analyses can lead to erroneous conclusions concerning the validity of the QSSA. A better approach is to consider

[3]Normal hyperbolicity of S_0 ensures the algebraic multiplicity of $D_1 f(x,0)$'s trivial eigenvalue is equal to its geometric multiplicity, k. It is possible in applications that the algebraic multiplicity is $k+1$ whereas the geometric multiplicity is k. In this case, S_0 will remain a k-dimensional submanifold of $\mathbb{R}^n$ since rank $D_1 f(x,0) = n - k$ along S_0, but normal hyperbolicity fails since one non-trivial eigenvalue of $D_1 f(x,0)$ will lie at the intersection of the real and imaginary axes.

the mass action equations for a general reaction network,

$$\dot{x} = f(x, p), \quad f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \tag{1.13}$$

where p denotes an m-tuple of parameters. Now suppose there is a point, p^*, in parameter space such that

1. The set $S_0 = \{x \in \mathbb{R}^n : f(x, p^*) = 0\}$ is a k-dimensional submanifold of $\mathbb{R}^n$ with $0 < k < n$, and

2. S_0 is normally hyperbolic;

then we call p^* a Tikhonov-Fenichel parameter value (TFPV) [65, 66]. Treating coordinates in parameter space as column vectors and expanding $f(x, p)$ near p^* in a direction ρ, $p = p^* + \varepsilon\rho$, yields[4]

$$\dot{x} = f(x, p^*) + \varepsilon D_2 f(x, p^*)\rho \tag{1.14}$$

which is in the form (1.12). Hence, due to the inherent multi-dimensionality of parameter spaces associated with chemical reaction networks, a path in parameter space must be specified along which to take the limit as $\varepsilon \to 0$ and $p \to p^*$.

As an example, consider the Michaelis–Menten system (1.2) again. The parameter space is a subset of $\mathbb{R}^4_{\geq 0}$ with $p = (e_0, k_1, k_{-1}, k_2)^T$. Note that $p^* := (0, k_1, k_{-1}, k_2)$ is a TFPV as long as k_1, k_{-1}, k_2 are bounded above and below by positive constants.

To put (1.2) into perturbation form, we specify the path in parameter space along which we approach p^*. Since p^* coincides with a trivial total enzyme concentration, we define the

[4]The operator D_2 denotes differentiation with respect to the second argument, p.

curve, $\ell(\cdot)$, parameterizable by ε, that passes through p^*

$$\ell(\varepsilon) = \begin{pmatrix} 0 \\ k_1 \\ k_{-1} \\ k_2 \end{pmatrix} + \varepsilon \begin{pmatrix} \widehat{e}_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = p^* + \varepsilon\rho \tag{1.15}$$

where $e_0 \mapsto \varepsilon\widehat{e}_0$, with ε denoting the small magnitude of e_0 and $\widehat{e}_0$ carrying the (arbitrary) units of concentration with $|\widehat{e}_0| = 1$. Thus, the perturbation form of (1.2) for small e_0 is

$$\begin{pmatrix} \dot{s} \\ \dot{c} \end{pmatrix} = \begin{pmatrix} (k_1 s + k_{-1})c \\ -(k_1 s + k_{-1} + k_2)c \end{pmatrix} + \varepsilon\widehat{e}_0 k_1 s \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{1.16}$$

When $p = p^*$, the mass action equations for s and c are

$$\dot{s} = \quad (k_1 s + k_{-1})c \tag{1.17a}$$

$$\dot{c} = -(k_1 s + k_{-1} + k_2)c \tag{1.17b}$$

and by inspection the critical manifold is the s-axis:

$$S_0 := \{(s, c) \in \mathbb{R}^2 : c = 0\}. \tag{1.18}$$

The Jacobian of (1.17) along S_0 is

$$D_1 f(s, c, 0) = \begin{pmatrix} 0 & k_1 s + k_{-1} \\ 0 & -k_1 s - k_{-1} - k_2 \end{pmatrix} \tag{1.19}$$

and has rank 1 for all $(s, c) \in S_0$. Moreover, since the non-trivial eigenvalue of the Jacobian,

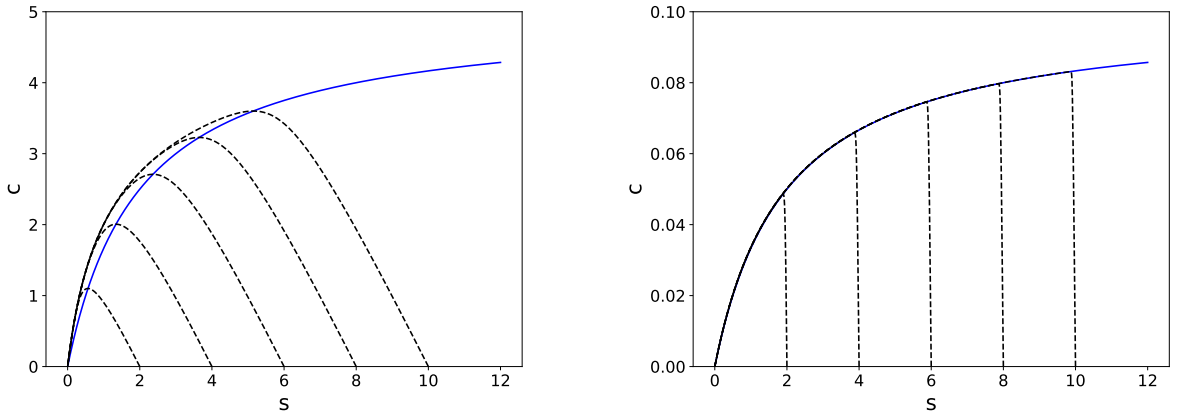


Figure 1.1: **The slow flow for the MM mechanism converges to the standard QSS approximation as $e_0 \to 0$.** The dashed black curves are the numerical solutions to the full MM ODE model (1.2) for different initial conditions. The solid blue curve is the QSS curve for the standard Fenichel reduction (1.24). **Left**: $e_0 = 5$. **Right**: $e_0 = 0.1$. **Both**: The initial conditions are $s_0 \in \{2, 4, 6, 8, 10\}$ and $c_0 = 0$. The parameter values are $k_1 = 1, k_{-1} = 1, k_2 = 1$.

λ_{MM}, is strictly less than zero,

$$\lambda_{MM} := -(k_1 s + k_{-1} + k_2),$$

and the critical manifold is attractive. Therefore S_0 is normally hyperbolic and attractive.

In order to compute the projection matrix Π^{S_0}, we assume the factorization $f(s, c, 0) = N(s, c)h(s, c)$. Since the rank of the Jacobian $\mathrm{D}_1 f(s, c, 0)$ is constant and equal to 1 for all points on the critical manifold, the factorization must hold locally [184]. Moreover, the zero-level set of $f(s, c, 0)$ corresponds to S_0, the columns of $N(s, c)$ form a basis for the range of $\mathrm{D}_1 f(s, c, 0)$ and the row vectors of $\mathrm{D}h(s, c)$[5] form a basis for the orthogonal component of $\ker \mathrm{D}_1 f(s, c, 0)$:

$$\text{column space } N(s, c) = \text{range } \mathrm{D}_1 f(s, c, 0), \quad \forall (s, c) \in S_0, \tag{1.20a}$$

$$\text{row space } \mathrm{D}h(s, c) = \ker^T \mathrm{D}_1 f(s, c, 0), \quad \forall (s, c) \in S_0. \tag{1.20b}$$

[5] $\mathrm{D}h(s, c)$ is the 2×2 Jacobian of $h(s, c)$ with respect to (s, c).

For the layer problem (1.17), it is straightforward to compute the factorization,

$$f(s,c,0) = N(s,c)h(s,c), \quad \text{with} \quad N(s,c) := \begin{pmatrix} k_1 s + k_{-1} \\ -(k_1 s + k_{-1} + k_2) \end{pmatrix}, \quad \text{and} \quad h(s,c) := c, \tag{1.21}$$

as well as the derivative of $h(s,c) = c$:

$$\mathrm{D}h(s,c) = (0 \;\; 1). \tag{1.22}$$

The oblique projection matrix onto the tangent space of S_0 is given by

$$\Pi^{S_0} := \mathbb{I}^{n\times n} - N(\mathrm{D}hN)^{-1}\mathrm{D}h = \begin{pmatrix} 1 & \mu(s) \\ 0 & 0 \end{pmatrix}, \qquad \mu(s) = \frac{K_S + s}{K_M + s} \tag{1.23}$$

and the projection of the second term of (1.16) onto TS_0 yields

$$\begin{pmatrix} \dot{s} \\ \dot{c} \end{pmatrix} = -\frac{\varepsilon \widehat{e}_0 k_2 s}{K_M + s} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv -\frac{e_0 k_2 s}{K_M + s} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{1.24}$$

The reduced system (1.24) is known as the standard quasi-steady-state approximation (sQSSA).[6] The interpretation behind the derivation (1.24) is that the flow on S_0^ε converges to (1.24) as $e_0 \to 0$ (or $\varepsilon \to 0$) with all other parameters held fixed and positive, see Figure 1.1 for numerical evidence. In fact, solutions to (1.2) starting on the s-axis will converge to (1.24) since the s-axis coincides with the critical manifold, S_0.

1.1.2 Quantitative and Qualitative Analysis of Reduced Models

Quasi-steady-state (QSS) approximations are extensively employed for real-world applications such as informing drug development strategies. Researchers commonly infer kinetic parameters by fitting experimental time-course data to QSS models through nonlinear re-

[6] In (1.24), we have transformed $\varepsilon\widehat{e}_0$ back to e_0 for clarity, and will continue to do this from this point forward.

gression techniques [158, 24]. These parameters offer critical insight into how effectively enzymes, substrates, inhibitors, or activators function within a biochemical pathway. Consequently, any reduction theory including geometric singular perturbation theory (GSPT) underpins the reliability and accuracy of these applications [168]. To ensure accurate parameter estimation, experiments must be designed to satisfy the assumptions underlying QSS approximations.

As a result, a key element of Fenichel theory is the in-depth understanding of the small perturbation parameter ε. As established through TFPV theory, the smallness of the dimensionless parameter ε is an indication of the accuracy of the associated reduced model. However, a persistent and under-explored question remains: "How *close* to a TFPV should a parameter value be for the QSS reduction to be suitably accurate?" Equivalently "How *small* should ε be for the QSS reduction to be suitably accurate?" Contrary to our expectations, terms such as *close*, *small* and *suitably accurate* have been defined more qualitatively than quantitatively.

For perturbation systems present in the standard form (1.6), with a clear distinction in fast and slow reactions, the small perturbation parameter ε is proportional to the ratio of the fast and slow timescales. More generally, the parameter emerges from good timescale estimates that utilize a heuristic scaling analysis [164, 126]. For instance, classical works, including Heineken et al. [71] and Segel [164], suggest some scaling variables for QSS justification in the Michaelis–Menten mechanism. However, as Eilertsen et al. [47] note, heuristic methods often rely on creative choices that can lead to inaccurate conclusions and fail to provide rigorous quantitative estimates for model validity. In fact, the qualifier for the standard QSSA (1.24) suggested by Segel and Slemrod [163] was later challenged and revised for a more restrictive parametric condition by Kang and Kurtz [93], Mastny et al. [117], Eilertsen et al. [48].

Several recent works address these gaps and propose algorithmic approaches to find estimates for small perturbation parameters. For reductions obtained through Fenichel and

TFPV theory, Goeke et al. [65, 66] determine conditions on perturbation parameters by measuring the discrepancy between the perturbed system and the reduced system. Palsson and Lightfoot [130] and Eilertsen et al. [49] present an eigenvalue approach using linearization of the system near equilibrium points. Here, local timescale estimates are obtained through the ratio of eigenvalues of the Jacobian near the critical manifold, though these calculations can become intractable in complex systems.

Despite these advances, challenges remain in applying reduction theory to practical applications. In multi-dimensional parameter spaces, different TFPVs may lie close together. Even when one reduction is justified, proximity to another TFPV may complicate its optimal applicability. Furthermore, many popular reduced models are not derived directly from GSPT or TFPV frameworks, making it difficult to extend existing quantitative estimates for ε. Techniques like timescale separation or eigenvalue ratios may fail when Fenichel or TFPV assumptions, such as normal hyperbolicity, break down. In such cases, the Jacobian may yield trivial eigenvalues, rendering ratio-based criteria undefined.

Although workarounds exist for finding reduced models in such scenarios, there is limited work on the mathematical justification of these approximations. This thesis contributes to the broader understanding of QSSA validity, including identification of suitable ε values and associated quantitative error estimates.

1.2 Mathematical Modeling for Sleep-Wake Regulation

Sleep-wake regulation represents a prominent domain in biological research where mathematical modeling has significantly advanced our understanding of neural dynamics. Healthy sleep patterns are essential for cognitive functioning, cardiovascular and metabolic regulation, and overall development in animals, including humans [165, 25, 85, 58]. These physiological benefits motivate the development of mathematical models capable of explaining and predicting the underlying mechanisms governing sleep.

Over the years, a wide spectrum of models has been developed to describe sleep-wake dy-

namics. These range from deterministic ODE models to statistical and probabilistic frameworks, each incorporating biophysically inspired parameters and processes [138, 34, 131]. Such models aim to replicate observed patterns of sleep transitions, duration, and architecture. A crucial next step in deepening our understanding is the accurate estimation of model parameters that reflect individual and species-specific variability in sleep timing and structure.

Parameter estimation in this context allows researchers to link measurable biological variables such as Electroencephalogram (EEG) and Electromyography (EMG) recordings, to the internal dynamics of sleep-regulating neuronal circuits. This facilitates comparative studies across species, developmental stages, and environmental factors [121, 166, 6]. Moreover, it opens doors to identifying physiological or pathological deviations in sleep regulation, with potential clinical applications.

In this section, we begin by providing an overview of the experimental data typically used in sleep-wake regulation research, including laboratory methods of scoring sleep stages. We then introduce the neurophysiological foundations of sleep control, particularly focusing on the flip-flop switch model that captures transitions between sleep and wakefulness using a minimal set of ODEs. Finally, we review key literature on parameter estimation in biological modeling and set the motivation for the methodological contributions of this work.

1.2.1 Neurobiology for Data Retrieval and Flip-Flop Models

Sleep-wake data is typically a set of discrete temporal recordings, divided among the states of wakefulness, Rapid Eye Movement (REM) sleep and Non-REM (NREM) sleep. These states are observed as a result of distinct electrical activity in the brain recorded through biomarkers such as EEG and EMG readings. The EEG and EMG signals are commonly analyzed using time-frequency methods, such as spectrograms for EEG, and are converted into sleep-wake stage hypnograms using fixed-duration resolution (for example, 2.5 seconds intervals in [131]). Variability in activity or alertness is identified through the power in different frequency bands.

In particular, EEG activity in the theta band (5 - 12 Hz) and sigma band (12 - 20 Hz) along with low EMG power relates to a NREM sleep state. This associates with light sleep along with minimal eye and muscle movements. A subtype of NREM sleep is further characterized with low frequency, high amplitude EEG activity in the delta band (0.5 - 4 Hz), referred to as the slow wake sleep (SWS) [131]. REM sleep is observed through low EMG power but high EEG frequency activity which may extend into the gamma band (100 - 150 Hz). Wake, on the other hand, relates to high EMG power and high frequency and low amplitude EEG signals. States with high gamma activity are usually assigned as wake if they were not already classified as REM [131].

In addition to EEG and EMG, other physiological signals such as electrooculogram (EOG) and electrocardiogram (ECG) are aMGABlso used to classify sleep states. Similar to EEG and EMG readings, these polysomnography experiments are highly accurate and can only be performed in labs. Another data collection method that has garnered research attention is actigraphy [15]. Actigraphy, while less precise than polysomnography, involves data collected from wearable devices that track movement and heart activity and is primarily used in ambulatory or home settings.

Within the past few decades, experimental studies have uncovered some key features of the neurophysiological mechanisms governing sleep-wake regulation, including the identification of wake-promoting and sleep-promoting neuronal populations in various regions of the brain [153]. Neuronal populations that drive arousal consist of monoaminergic neurons.[7] These neurons are mostly present in the brainstem and hypothalamus region including the locus coeruleus (LC), dorsal raphe (DR), tuberomammillary nucleus (TMN), cholinergic pedunculopontine and laterodorsal tegmental nuclei (PPT/LDT) and the basal forebrain (BF) [88, 157]. During periods of heightened activity, neurons in these regions promote wakefulness by sending signals to higher cortical areas using both dorsal and ventral pathways. The wake-promoting neurons are associated with the neurotransmitter noradrenaline

[7]Monoaminergic neurons are the neurons that use monoamine neurotransmitters.



Figure 1.2: **The activity in presynaptic neuronal populations govern neurotransmitter concentrations which affect the activity in postsynaptic neuronal populations.**

(NE), among others such as serotonin and histamine, that drive their interaction with sleep-promoting neurons.

Sleep, on the other hand, is promoted by the neurons in ventrolateral (VLPO) and median preoptic nuclei (MnPO) regions. Various studies reveal the role of gamma-amino butyric acid (GABA), the neurotransmitter released by sleep-promoting neurons, in sleep activation and timing [112, 57, 179, 181].

One of the simplest sleep-wake models is the two-state flip-flop ODE model that relies on the mutual inhibition between wake-promoting and sleep-promoting populations [154]. The reciprocal inhibition is controlled by the synaptic connection established through the neurotransmitters released by both populations. These models use a firing-rate-based formalism, representing the activity of wake and sleep neuronal groups using firing rates, $f_W(t)$ and $f_S(t)$ respectively, as functions of time. The concentrations of the neurotransmitters can vary over time or admit a steady state value. In this framework, the firing rate of a presynaptic population drives neurotransmitter release, which in turn modulates the activity of postsynaptic populations, refer to Figure 1.2.

Sleep-wake regulation is further governed by two additional biological processes: the circadian rhythm from the suprachiasmatic nucleus (SCN) that drives the (roughly) 24 hour sleep cycle in most animals and the homeostatic sleep drive that dictates the switch between wake and sleep states mimicking the accumulation of fatigue over longer wake periods [18, 33].

In Chapter 6, we focus on including the homeostatic sleep drive as an input to the simple two-state flip-flop model [2]. The only known biomarker for sleep homeostasis is the power

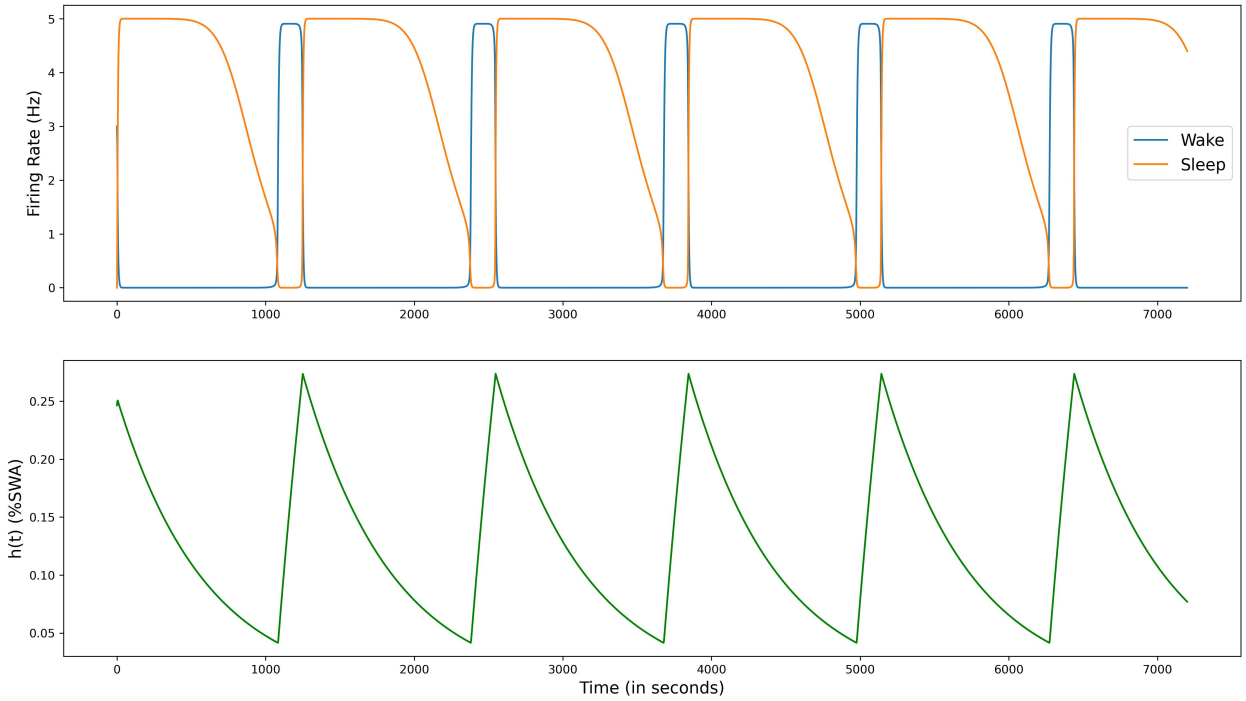


Figure 1.3: **Time traces showing the variation of f_W, f_S and $h(t)$ over 2 hours**. The traces are obtained by numerically solving a deterministic ODE flip-flop system with initial conditions $f_W(0) = 3$, $f_S(0) = 0$, $h(0) = 0.24637$. The system is solved for every 1.25 seconds. All the details for the model are presented in Chapter 6.

in the EEG slow wave frequency band (slow wave activity, SWA) which suggests exponential accumulation and dissipation [53]. The buildup of adenosine, a molecule that promotes sleep pressure and is blocked by caffeine, is shown to be a key factor contributing to increased slow-wave activity [4]. In the model, the homeostatic drive, denoted as $h(t)$, accumulates during periods of wakefulness due to the activation of the wake-promoting population. As $h(t)$ increases, it begins to stimulate the sleep-promoting population, initiating a transition to the sleep state. In the sleep phase, diminished wake-promoting activity causes a gradual reduction in $h(t)$, which eventually suppresses the sleep-promoting population, leading to a reactivation of the wake state. Figure 1.3 depicts an example time trace for the flip-flop nature of $f_W(t), f_S(t)$ and $h(t)$.

Beyond the flip-flop model, a variety of more detailed and biologically realistic models have been proposed to simulate dynamics of wake, NREM and REM sleep transitions.

These include multi-population models that incorporate finer neuroanatomical detail and diverse neurotransmitter systems [146], network-based models of cortical and thalamic interactions [8], and spiking neuron models that represent membrane potential dynamics with high temporal precision [135]. Probabilistic and state-space models have also been employed to capture sleep as a stochastic process with hidden latent variables driving transitions [136, 13, 131].

While these models offer rich insights into neuronal mechanisms, they are often computationally intensive and harder to interpret or optimize in data-driven contexts. The overall utility of mathematical models is gained through their accuracy in simulating real-world sleep-wake data. This is achieved by solving complex inverse problems and estimating the physiological parameters. In contrast, the two-state flip-flop model, despite its simplicity, retains key qualitative features of sleep-wake switching and offers significant advantages in computational efficiency and parameter interpretability. Thus, in Chapter 6, we modify a two-state model to account for the randomness and irregularities in experimental rodent sleep-wake data, making it more suitable for parameter optimization.

1.2.2 Methods and Measures for Parameter Estimation

Estimating biologically relevant parameters from mathematical models is pivotal in understanding complex biological systems. Some widely adopted data-driven optimization methods are least-squares estimation (LSE), maximum likelihood estimation (MLE) and maximum a posteriori estimation (MAP) [14]. While the former relies on direct comparison of true versus predicted data without any distributional assumptions, the latter techniques aim to maximize the likelihood of observed data given specific parameter values [114]. Among these, MLE assumes a uniform prior whereas MAP incorporates the specific prior distribution of parameters, enhancing estimation accuracy when such information is available. However, making an informed guess for the priors of biophysical parameters remains challenging. Additionally, the non-linear and intricate nature of sleep-wake regulation models often leads to other challenges like local maxima in the likelihood landscape, complicating the optimization

process.

To address the optimization challenges, Bayesian inference methods, particularly Markov Chain Monte Carlo (MCMC) techniques, have gained prominence [178]. MCMC methods sample parameter values from prior distributions and update them based on observed data to approximate posterior distributions. Alternative approaches include spline approximations and smoothing methods that offer a way to estimate parameters by smoothing data and reducing computational complexity [144, 188]. Similarly, the WENDy (Weak form Estimation of Nonlinear Dynamics) framework provides a means to estimate parameters in systems with inherent noise, enhancing model reliability [19].

Recent advancements in optimization algorithms, particularly in the realm of deep learning, have opened new avenues for parameter estimation. Neural networks (NNs) have demonstrated success in various fields, including biology. In the context of biological systems, NNs have been employed to solve both forward and inverse problems [156]. Work by Dua [35] incorporates artificial neural networks that learn the solutions to a system of ODEs and estimate parameters by minimizing the error between the actual vector field and the NN-generated vector field. Gonçalves et al. [68] and Rudi et al. [149] assume a prior density function and learn posterior density estimates for the parameters using neural networks. Moreover, Generative Adversarial Networks (GANs) have been utilized to infer parameters by generating synthetic data that closely resembles real observations, again utilizing a prior and estimating a posterior for the parameter values [151]. While these methods offer improved accuracy over traditional techniques like MCMC, they often come with increased computational costs.

In the realm of sleep-wake models, particularly in modeling rodent sleep, the irregular occurrence of sleep states poses unique challenges. Implementing neural network-based models with complex ODE models can become computationally intensive, and acquiring precise priors or likelihoods for physiological parameters is often unfeasible. To navigate these obstacles, we propose developing a robust framework that leverages gradient descent algorithms,

a cornerstone of deep learning, for parameter estimation in sleep-wake regulation models. In Chapter 6, we delve into methods for introducing exogenous noise into ODE models, facilitating automatic differentiation for efficient gradient computation. Moreover, determining an appropriate error metric that accurately captures discrepancies between model outputs and empirical data is crucial, especially when temporal alignments are imperfect. Most of the previous works rely on a direct comparison of model outputs and datasets, often resorting to sum of squares errors or absolute errors. Chapter 6 provides a comprehensive discussion on error metrics tailored for noisy model outputs.

1.3 Overview and Contributions

In this section, we provide a brief overview of the thesis including the main aims and contributions of all chapters. The appendices include extra information for the chapters and are presented after the chapters. All the references are included at the end of the thesis. Broadly put, Chapters 2-5 focus on the analysis of quasi-steady-state approximations in models of enzyme kinetics whereas Chapter 6 addresses parameter estimation for a model of sleep-wake regulation in rodents.

Chapter 2 - Stochastic Enzyme Kinetics and the Quasi-Steady-State Reductions: Application of the Slow Scale Linear Noise Approximation à la Fenichel

The linear noise approximation models the random fluctuations from the mean-field model of a chemical reaction that unfolds near the thermodynamic limit. Specifically, the fluctuations obey a linear Langevin equation up to order $\Omega^{-1/2}$, where Ω is the volume of the chemical system. In the presence of disparate timescales, the linear noise approximation admits a quasi-steady-state reduction referred to as the *slow scale* linear noise approximation (ssLNA). Curiously, the ssLNAs reported by past works such as those by [129, 170, 74] are slightly different. The differences in the reported ssLNAs motivate the work in this chapter.

Specifically, we derive the ssLNA directly from geometric singular perturbation theory and explain the origin of the different ssLNAs in the literature. We find that the ssLNA derived

from GSPT is closer to the ssLNA(s) established in the literature when the perturbation system is present in standard form (1.6) [48].

Moreover, we extend the ssLNA derived from geometric singular perturbation theory to a non-classical singularly perturbed problem consisting of a transcritical bifurcation point that results in the loss of normal hyperbolicity. To the best of our knowledge, this is the first extension of the ssLNA to a non-hyperbolic singularity. In so doing, we disprove the commonly-accepted Segel-Slemrod qualifier,

$$e_0 \ll K_M + s_0, \tag{1.25}$$

for the validity of stochastic QSS approximation of the Michaelis–Menten reaction mechanism (1.2) [163, 152]. We use insights from the derived ssLNA and Gillespie algorithm to show that a more restrictive condition,

$$e_0 \ll K_M, \tag{1.26}$$

is necessary to ensure the validity of the sQSSA in the stochastic regime, in agreement with other stochastic works by Kang and Kurtz [93] and Mastny et al. [117]. This result also upholds the widely-known Reich-Sel'kov qualifier, $e_0 \ll K_M$, for the deterministic sQSSA for the MM mechanism [145, 42], bridging the application of GSPT between the deterministic and stochastic regimes.

Chapter 3 - Accuracy versus Predominance: Reassessing the Validity of the Quasi-Steady-State Approximation

The application of the standard QSS approximation to the Michaelis–Menten reaction mechanism (1.3) is a textbook example of biochemical model reduction, derived using singular perturbation theory. However, determining the specific biochemical conditions that dictate the validity of the standard quasi-steady-state approximation remains a challenging endeavor. Long standing research suggests that the accuracy of the standard quasi-steady-state approximation improves as the ratio of the initial enzyme concentration, e_0, to the

Michaelis constant, K_M, decreases [145, 48]. The parametric qualifier, $e_0/K_M \ll 1$, referred to as the Reich-Sel'kov condition, is also analysed in Chapter 2 in the stochastic regime. In this chapter, we examine this ratio further and its implications for the accuracy and validity of the standard quasi-steady-state approximation in the deterministic model as compared to other quasi-steady-state reductions in its proximity.

First, using the theory of fences and anti-funnels, we precisely locate the slow manifold (unique solution) in the (s, c) phase plane for the MM mechanism. This provides sharp error estimates on the long-term accuracy of the sQSSA, improving on past findings [51, 125]. We also recover the dependency of the sQSSA on the Reich-Sel'kov condition without relying on any heuristic scaling argument.

Second and more importantly, we show that while e_0/K_M provides an indication of the sQSSA's accuracy, its predominance relies on a small ratio of e_0 to the Van Slyke-Cullen constant, K. Here, we define the *predominance* of a quasi-steady-state reduction when it offers the highest approximation accuracy among other well-known reductions with overlapping validity conditions. This is the first time that the notion of the validity of the QSS approximations is linked to both its accuracy as well as its predominance. When the validity conditions for multiple reduced models lie in close proximity in the parameter space, this distinction ensures selection of the most appropriate reduction for the most optimal experimental design. To this end, for the MM mechanism, we conclude that the more restricted qualifier

$$e_0 \ll K \tag{1.27}$$

offers the most accurate measure of the validity of the standard quasi-steady-state approximation [167].

Chapter 4 - Gaining Hyperbolicity: A Blow-up Approach to Understanding Quasi-Steady-State Approximations

Given the pervasive use of quasi-steady-state reductions, extensive research has sought to extend GSPT and TFPV theories to more realistic, non-idealized settings. A central limitation of these theories is their reliance on a normally hyperbolic critical manifold, a condition often violated in real biochemical systems. Many mechanisms feature bifurcation points, center manifolds, folds, canards, or oscillations; phenomena typically associated with the loss of normal hyperbolicity [104, 105, 38].

While past work by Krupa and Szmolyan [103] has extended GSPT to accommodate non-hyperbolic singularities, these developments are largely confined to systems expressed in a canonical or normal form. There remains limited analysis of QSS reductions in real-world chemical systems which do not conform to normal form bifurcation theory. In this chapter, we address this gap by studying the Intermolecular Autocatalytic Zymogen Activation (IAZA) mechanism under a perturbation scenario which exhibits a dynamic transcritical bifurcation point. Unlike systems treated in existing theory, the IAZA model does not naturally fall into a normal form structure.

To this end, we find that the dissociation parameter k_{-1} controls the extent to which the dynamics align with or deviate from transcritical normal form behavior. The irreversible IAZA system (when $k_{-1} = 0$) represents a degenerate case in the sense of Krupa and Szmolyan [103], where trajectories might exhibit faux canard-like solutions near the bifurcation. As k_{-1} increases, the dynamics deviate from those predicted by normal form theory, before eventually entering a completely different and normally hyperbolic regime. A novel contribution of this chapter is the identification of the dynamic imperfection induced by the dissociation parameter, resulting in behavior resembling that of an imperfect bifurcation.

Using these structural insights, we also rigorously analyse the validity of two well-studied QSS reductions for the IAZA mechanism. First, a heuristic modification to Fenichel reduction is adopted by ignoring a small neighborhood of the bifurcation point. Although the accuracy

of this reduction has been shown to be $\mathbb{O}(\sqrt{\varepsilon})$ near the bifurcation point, the lack of a precise estimate for ε presents further challenges [46]. An alternative reduction - though not derived from GSPT - assumes that the complex concentration closely follows its nullcline (steady state). Despite its empirical accuracy, this nullcline-based reduction lacks rigorous mathematical justification, particularly near the bifurcation point [187]. We aim to resolve this predicament in this chapter.

Building on the blow-up method introduced by Krupa and Szmolyan [104], we analyze the IAZA system in a transformed (blown-up) coordinate space near the bifurcation. This transformation restores sufficient hyperbolicity, enabling analytical insight at the otherwise non-hyperbolic singularity. In the blown-up phase plane at the bifurcation point, we derive an explicit (analytic) solution to the desingularized system and consequently estimate the point of maximum complex concentration along the trajectory. To our knowledge, this is the first analytical estimate of the peak point, a key quantity for determining limiting reaction rates. We find candidate parametric qualifiers that validate the accuracy of the modified Fenichel reduction. The analytical solution also helps us understand the exceptional and unreasonably universal approximation accuracy of the nullcline reduction, especially when compared to the modified Fenichel reduction [50].

Collectively, these results advance the understanding of the IAZA phase plane, fill gaps in previous studies such as [46], and provide a generalizable framework for applying the blow-up method to chemical reaction mechanisms.

Chapter 5 - Gaining Space, not Hyperbolicity: A Blow-up Approach to Understanding Quasi-Steady-State Approximations

Another limitation of geometric singular perturbation theory (GSPT) pertains to its applicability to systems that do not conform to the Tikhonov standard form (1.6). The foundational Tikhonov and Fenichel theories provide a robust framework for systematic reduction of two-timescale chemical mechanisms. However, for systems that are not present in standard form, the fast-slow dynamic structures are not very apparent and the convergence of

perturbed flow to the QSS approximation is not well understood. Previous works show that, unless the solutions start close to the critical manifold, we can explain the convergence by finding a suitable transformation of systems in non-standard form to systems in standard form. However, this approach often relies on heuristic transformations that may not generalize well and can become intractable even for relatively simple mechanisms.

In this chapter, we propose a systematic approach to uncover fast-slow dynamics in systems that do not adhere to the standard form. In particular, we introduce a novel application of the blow-up method to expand the phase space near the critical manifold, enabling a detailed analysis of trajectory behavior. Unlike the use of blow-up in Chapter 4, where the goal was to gain normal hyperbolicity, our application focuses on revealing latent fast-slow structures by gaining space in transformed phase planes.

We illustrate this approach with the two classic enzyme kinetics models studied in the thesis: the Intermolecular Autocatalytic Zymogen Activation (IAZA) and Michaelis–Menten (MM) mechanisms. Despite lacking an explicit standard-form structure, both systems exhibit behavior typical of such systems: a rapid accumulation of the fast species (intermediate complex) while the slow species (zymogen or substrate) remains nearly constant. This separation becomes more pronounced in the singular limit as shown in Figure 1.1, yet the underlying justification for this behavior has remained unclear.

Our analysis shows that, in the blown-up phase space, these systems do exhibit standard-form dynamics, with a clear distinction between transformed fast and slow variables. The critical manifolds of the local (blown-up) layer problem coincide with the QSS curves, and we demonstrate that trajectories - both near and away from the critical manifold - converge to these QSS curves. This result explains the observed universal convergence of solution trajectories to QSS curves, regardless of initial conditions, and enhances the theoretical understanding of QSS reductions in enzyme kinetics. A key contribution of this chapter is a generalizable blow-up algorithm that aims to reveal timescale separation in fast-slow systems, aiding the extension of GSPT to real-world settings.

Chapter 6 - Gradient-based Optimization of Noisy ODE Models for Polyphasic Rodent Sleep-Wake Behavior

Understanding the regulation of sleep-wake states in mammals, particularly rodents, is a central topic of research in computational neuroscience. Mathematical models have been instrumental in describing transitions between wake, REM sleep, and NREM sleep offering insights into underlying biophysical mechanisms [2, 17, 16]. Unlike humans, rodents exhibit poly-phasic and fragmented sleep patterns, characterized by frequent short wake bouts interspersed with REM and NREM sleep bouts. This complicates modeling efforts and demands approaches that can capture both deterministic and stochastic features of sleep patterns.

Even though ODE-based models have been the primary focus of theoretical advancements, previous data-driven works instead rely on probabilistic models that capture the variability observed in experimental data through sleep state transitions and fluctuations [13, 131]. An alternative but understudied strategy is the addition of noise into ODE-based models [34]. However, estimating biologically meaningful parameters in these noisy ODE models remains difficult. While Bayesian and surrogate modeling approaches have gained popularity, gradient-based optimization, despite being a cornerstone in parameter inference, has seen limited use in this context [90, 68, 178].

In this chapter, we present a framework for fitting noisy, mechanistic ODE models to discrete sleep-score data using gradient-based optimization. We adapt a physiologically motivated model of interacting neuronal populations, incorporating exogenous noise in the firing of wake-promoting neurons to produce sleep-wake patterning consistent with rodent sleep recordings [155, 183].

The framework leverages automatic differentiation for efficient gradient computation and parameter estimation. We also propose additional methodology to ensure robust parameter optimization including efficient error metrics, parameter sensitivity analysis, custom performance evaluation measures aligned with sleep classification, and data augmentation techniques.

This chapter contributes to the fitting of a simple, physiologically motivated ODE model in order to capture the fine temporal architecture of rodent sleep data, elucidating the necessary modeling choices and their biological implications. While focused on sleep-wake modeling, our methods generalize to other domains requiring parameter inference in noisy ODE systems without relying on priors or likelihoods.

CHAPTER 2

Stochastic Enzyme Kinetics and the Quasi-Steady-State Reductions: Application of the Slow Scale Linear Noise Approximation à la Fenichel

2.1 Introduction

As established, the reduction of deterministic models is generally achieved through the application of Tikhonov's theorem [173] and Fenichel theory [54, 55]. Several analyses of enzyme-catalyzed reactions have made good use of singular perturbation theory to generate approximations referred to as *quasi-steady-state* (QSS) approximations or reductions [44, 42]. In fact, over the last decade, much progress has been made in developing and applying the formalism of Fenichel theory to chemical kinetics, and the culmination of the recent literature has turned up some surprising results. First and foremost, the advent of Tikhonov-Fenichel parameter value (TFPV) theory, developed extensively by Goeke et al. [66, 65], has rigorously demonstrated that not all QSS reductions emerge as a result of a singular perturbation scenario, despite what scaling and numerical simulations might suggest [127, 45]. TFPV theory has also enhanced our understanding of the singular perturbation structure (when applicable) to pertinent reaction models, which has led to the discovery of bifurcations and other interesting phenomena present in the singular vector fields of the model equations [42]. Most surprising, however, is the revelation that traditional scaling methods may lead to erroneous conclusions concerning the mathematical origin and justification of QSS reduction

in chemical kinetics (see, for example, Goeke et al. [64], Section 4, as well as Eilertsen et al. [45]).

Given the recent developments in the deterministic theory of QSS reduction, the natural question to ask is: *Do any of these developments have something important to say about model reduction in the stochastic realm?* Model reduction is more challenging in the stochastic regime, but rigorous reduction methods that leverage the presence of disparate timescales do exist (see, for example, [92, 101, 93]). The focus of this chapter is on the application of QSS reduction in the linear noise regime, where stochastic fluctuations from the deterministic mean-field model are governed by a linear Langevin equation called the *linear noise approximation* (LNA). The general methodology for QSS reduction in the LNA regime, called the *slow scale linear noise approximation* (ssLNA), is by Thomas et al. [172], Pahlajani et al. [129], and Herath and Del Vecchio [74]. Interestingly, the reported ssLNAs are slightly different, and this raises the question: Where do these differences come from, and are they critical? The intent of this chapter is three-fold: (i) to explain why different ssLNAs exist in the literature, (ii) to introduce recent developments of deterministic QSS theory to the stochastic community, and (iii) to demonstrate techniques to extend the ssLNA to specific non-classical singularly-perturbed problems. In what follows, we derive the ssLNA directly from GSPT and demonstrate where the differences emerge between the GSPT-derived ssLNA and the ssLNAs of Thomas et al. [172], Pahlajani et al. [129], and Herath and Del Vecchio [74]. We also discuss the role of the GSPT-derived ssLNA in the QSS reduction of the chemical master equation (CME) and use it to debunk a commonly used validity criterion in the literature.

2.2 Quasi-Steady-State Approximations of Michaelis–Menten Mechanism

The Michaelis–Menten mechanism (1.2) serves as a benchmark enzymatic reaction for the extensive study of quasi-steady-state approximations and their validity. In Chapter 1, we

introduced the mechanism and presented the steps to derive its commonly used standard quasi-steady-state reduction. In this section, we quickly review the sQSSA and introduce the reverse QSSA for the MM system using Fenichel theory. Both these reductions will be referenced repetitively in the thesis.

2.2.1 Standard Quasi-Steady-State Approximation

When e_0 is small, the QSS approximation for the subsrate depletion over time is

$$\dot{s} = -\frac{k_2 e_0 s}{K_M + s}, \tag{2.1}$$

referred to as the standard quasi-steady-state approximation. See Chapter 1 for the detailed derivation. It also corresponds to the QSS curve for the steady state of complex concentration:

$$c = -\frac{e_0 s}{K_M + s}. \tag{2.2}$$

Note that the sQSSA (2.1) is not the result of singular perturbation problem that is in standard form (5.1). This is contrary to the justification established from scaling analyses that utilize non-dimensionalization (see, Heineken et al. [71], Segel and Slemrod [163]).

2.2.2 Reverse Quasi-Steady-State Approximation

The reverse QSSA (rQSSA) was originally defined by Segel and Slemrod [163] as a perturbation problem, and later investigated in detail by Schnell and Maini [161]. To preface the derivation of the rQSSA as a Fenichel reduction, we remark that there are two common conditions that emerge in biochemical applications:

1. $D_1 f(x, 0)$ changes rank in at least one point belonging to the critical set. This happens, for example, if $D_1 f(x, 0) = 0$ (a null matrix) at some point belonging to the set $f(x, 0) = 0$. See definition 1.2

2. The zero eigenvalue of the Jacobian evaluated at at least one point in S_0 has an

algebraic multiplicity that is greater than the geometric multiplicity (the splitting (1.11) does not hold at such points).

The rQSSA is valid for small k_{-1} *and* small k_2, and is of the variety 1. In perturbation form this corresponds to

$$\dot{s} = -k_1(e_0 - c)s + \varepsilon\widehat{k}_{-1}c, \tag{2.3a}$$

$$\dot{c} = k_1(e_0 - c)s - \varepsilon(\widehat{k}_{-1} + \widehat{k}_2)c. \tag{2.3b}$$

The critical set is given by,

$$S_0 := \{(s,c) \in \mathbb{R}^2_{\geq 0} : c = e_0, 0 \leq s \leq s_0\} \cup \{(s,c) \in \mathbb{R}^2_{\geq 0} : s = 0, 0 \leq c \leq e_0\}. \tag{2.4}$$

The rank of the Jacobian of the layer problem along S_0 is not constant

$$\begin{aligned} &\text{rank } \ D_1 f(s,c,0) = 0, \ \text{ if } \ (s,c) = (0,e_0), \\ &\text{rank } \ D_1 f(s,c,0) = 1, \ \text{ otherwise,} \end{aligned}$$

and thus TFPV theory does not apply.[1] However, observe that the compact submanifolds

$$\begin{aligned} S_a^r &:= \{(s,c) \in \mathbb{R}^2_{\geq 0} : c = e_0, \ \ \varrho \leq s \leq s_0\}, \qquad 0 < \varrho, \\ S_a^b &:= \{(s,c) \in \mathbb{R}^2_{\geq 0} : s = 0, \ \ 0 \leq c \leq e_0 - \kappa\}, \quad 0 < \kappa < e_0 \end{aligned}$$

are normally hyperbolic and attracting. When $s(0) > 0$, trajectories will initially approach and follow S_a^r before eventually following S_a^b. In fact, a quick analysis reveals the existence of a transcritical bifurcation (see FIGURE 2.1).

By the projection methods above, it is straightforward to show that the QSS reductions

[1] The point $\pi^* = (e_0, k_1, 0, 0)^T$ is not a TFPV.

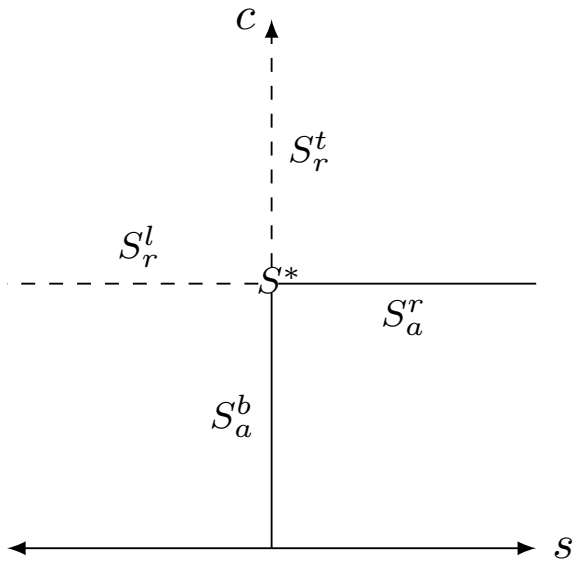


Figure 2.1: **A dynamic transcritical bifurcation occurs in the singular limit corresponding to $k_2 = k_{-1} = 0$ of the MM reaction mechanism.** At $S^\star$ the Jacobian has a double-zero eigenvalue. Along the dashed lines the Jacobian has one zero eigenvalue and one positive eigenvalue. Along the solid lines the Jacobian has one zero eigenvalue and one negative eigenvalue. At the bifurcation point $S^\star = (0, e_0)$, the lines $S_a^r \cup S_r^l$ and $S_r^t \cup S_a^b$ intersect and exchange stability.

obtained via projection onto $T_z S_a^r$ and $T_z S_a^b$ are, respectively:

$$\dot{p} = k_2 e_0, \qquad 0 \leq p < s_0 - e_0 \tag{2.7a}$$

$$\dot{p} = k_2(s_0 - p), \qquad s_0 - e_0 < p \leq s_0. \tag{2.7b}$$

As a concluding remark, note that we have successfully computed QSS reductions without a priori scaling and non-dimensionalization of the mass action equations. The ability to compute QSS reductions directly from the dimensional equations is a result of the TFPV theory developed by Goeke et al. [65, 66], which we have utilized here.

2.3 Stochastic Chemical Kinetics: Expansions, Reductions, and Approximations

In this section, we discuss QSS reduction in the stochastic regime. We introduce the chemical master equation (CME) and the derivation of the LNA via the Ω–expansion. We conclude with a review of the ssLNA as derived by Thomas et al. [172] and Pahlajani et al. [129], and we compare it to the GSPT-derived ssLNA.

2.3.1 Stochastic Chemical Kinetics Far from the Thermodynamic Limit: The Master Equation

Under physical conditions, a reaction occurs within a bounded volume, Ω. If the number of molecules in the system is finite, the reaction will *always* exhibit fluctuations. In fact, in the presence of random fluctuations and intrinsic noise, stochastic models provide a more physically realistic description of the kinetics when a system is far from the thermodynamic limit since the time interval between successive reactions becomes a random variable. The appropriate mathematical model depends on how "close" the system is to the thermodynamic limit.

If the chemical reaction consists of "R" elementary reactions, and the mixture is homogeneous and not diffusion limited then, far from the thermodynamic limit, the probability $P(Z,t)$ of finding the system in state Z at time t can be obtained from the solution to the CME (see, [62] and [91] for details),

$$\frac{\partial P(Z,t)}{\partial t} = \sum_{j=1}^{R} a_j(Z-\nu_j)P(Z-\nu_j,t) - a_j(Z)P(Z,t), \tag{2.8}$$

where ν_j are the stoichiometric vectors that correspond to the jth elementary reaction. If the system is in state Z when the jth reaction occurs, then the new state of the system will be $Z+\nu_j$. The functions a_j are called *propensity functions*. Dynamically the state of the system at time t is Z, and it moves from state Z to the state $Z+\nu_j$ within the infinitesimal

window $[t, t + \mathrm{d}t)$ with the probability

$$P(Z + \nu_j, t + \mathrm{d}t|Z, t) = a_j(Z)\mathrm{d}t. \tag{2.9}$$

The conditional probability (2.9) of jumping into the state $Z+\nu_j$ depends only on the present state of the system, which is called the Markov property.

The CME is possibly the most fundamental description of a chemical reaction. The difficulty is that closed-form solutions are rarely attainable. This begs question: Is it possible to derive physical models that exhibit stochasticity, but are nevertheless easier to analyze? The answer is yes, but the cost is that simplified models are usually only valid in monostable systems near the thermodynamic limit. The LNA is of this variety.

2.3.2 Approaching the Thermodynamic Limit: the LNA

To introduce the LNA, it is helpful to express the mass action equations in the form

$$\dot{z} = \mathcal{S}q, \quad F := \mathrm{diag(q)} \tag{2.10}$$

where $\mathcal{S}$ is the stoichiometric matrix. Additionally, we define the matrix F with the main diagonal q such that the diagonal components correspond to the elementary reactions of the chemical system. For example, the MM reaction mechanism (3.1) consists of three elementary reactions: the formation of complex, the disassociation of complex into S and E, and the disassociation of complex into E and P. Hence, the mass action system in form (2.10) is

$$\begin{pmatrix} \dot{s} \\ \dot{c} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} k_1(e_0 - c)s \\ k_{-1}c \\ k_2 c \end{pmatrix} := \mathcal{S}q. \tag{2.11}$$

To formally derive the LNA, one starts with the *operator form* of the CME,

$$\frac{\partial P(Z,t)}{\partial t} = \Omega \sum_{j=1}^{R} \left(\prod_{i=1}^{m} \mathbb{E}^{\mathcal{S}_{ij}} - 1 \right) a_j(z) P(Z,t), \tag{2.12}$$

where $z = Z/\Omega$ and $\mathbb{E}^{\mathcal{S}_{ij}}$ is the step operator:[2]

$$\mathbb{E}^{-\mathcal{S}_{ij}} a(Z) = a(\mathbb{E}^{-\mathcal{S}_{ij}} Z) = a(Z - \mathcal{S}_{ij} e_i). \tag{2.13}$$

Inserting the ansatz $Z = \Omega z + \Omega^{1/2} X$ into (2.12) and expanding (2.12) in powers of Ω yields (2.10) at zeroth-order in Ω. Thus, the mean of the stochastic trajectory obeys the mass action equation (2.10).

At order $\Omega^{-1/2}$, the equation that determines the randomly fluctuating departure from the mean, X, is a linear stochastic differential equation (SDE),

$$\mathrm{d}X = JX \, \mathrm{d}t + \Omega^{-1/2} \mathcal{S} \sqrt{F} \, \mathrm{d}W, \tag{2.14}$$

where J, the Jacobian, is $J := D(Sq)$, and W is a Wiener process. Collectively, (2.14) and the mass action equations comprise the LNA. On occasion we will express the LNA in the form

$$\dot{X} = JX + \Omega^{-1/2} \mathcal{S} \sqrt{F} \, \zeta(t), \tag{2.15}$$

where the Gaussian white noise, $\zeta(t)$, is understood to be the generalized derivative of W.

The LNA is notably simpler than the CME, since the Langevin equation (2.14) is *linear*, and the integration of linear stochastic differential equations of the form (2.14) is well-understood. The Fokker-Planck equation associated with (2.14) is also linear,

$$\frac{\partial \rho(X,t)}{\partial t} = \left(- \frac{\partial}{\partial X_i} (JX)_i + \frac{1}{2\Omega} D_{ij} \frac{\partial}{\partial X_i} \frac{\partial}{\partial X_j} \right) \rho(X,t), \tag{2.16}$$

[2]Here, e_i is the standard basis vector in $\mathbb{R}^n$.

where the diffusion matrix, D, is given by $D = \mathcal{S}F\mathcal{S}^T$.

As mentioned in the earlier sections, the reduction of the LNA based on timescale separation is the ssLNA developed by Thomas et al. [172, 171] and Pahlajani et al. [129]. In the nonlinear regime, Katzenberger [97] addressed reduction of SDEs of the form

$$\mathrm{d}x = w(x)\mathrm{d}t + \varepsilon G(x,\varepsilon)\mathrm{d}t + \sqrt{\nu}B(x)\ \mathrm{d}W \tag{2.17}$$

where ε and ν are extremely small (i.e., $\varepsilon, \nu \ll 1$). In short, Katzenberger [97] proved that provided specific conditions hold, SDEs of the form (2.17) converge, in a certain sense, to the reduced SDE,

$$\mathrm{d}x = \varepsilon\Pi G(x,\varepsilon)\mathrm{d}t + \sqrt{\nu}\Pi B(x)\ \mathrm{d}W + \nu\mathfrak{D}(x,\varepsilon), \tag{2.18}$$

where Π is a projection operator that maps to the tangent space of the critical manifold $S_0 := \{x \in \mathbb{R}^n : w(x) = 0\}$ that emerges when $\varepsilon = \nu = 0$, and $\nu\mathfrak{D}(x)$ is a noise-induced drift term. Parsons and Rogers [132] derived the explicit construction of Π and $\mathfrak{D}(x)$ in their analysis of fully nonlinear Langevin equations. Notably, Parsons and Rogers [132] did not discuss the reduction of noisy systems in standard form, and a projection operator Π that is consistent with Fenichel theory has not been defined for standard-form singularly perturbed systems in the linear noise regime. Such is the subject of the subsection that follows.

2.3.3 Projecting onto the Tangent Space of the Critical Manifold

Classical singular perturbation reduction of a deterministic system requires the existence of a normally hyperbolic critical manifold in the singular limit; this fact is non-negotiable. The reduction of the LNA is also straightforward, provided one has a well-defined critical manifold. The key observation in the LNA regime is to recognize that the dimension of the problem increases, but that the LNA is still of the form (2.17), and therefore the results of Katzenberger [97] are applicable. All that remains is to identify a normally hyperbolic critical manifold, its tangent space, and construct the unique projection operator, Π.

The Jacobian, $Dw(z)$, is equal to ND_μ if $z \in S_0$, and in standard form, the general LNA

is

$$\dot{z} = w(z) + \varepsilon G(z, \varepsilon), \tag{2.19a}$$

$$\dot{X} = ND\mu \cdot X + \varepsilon DG(z, \varepsilon) \cdot X \ + \Omega^{-1/2} \mathcal{S}\sqrt{F} \cdot \Gamma, \tag{2.19b}$$

where $\Gamma := (\zeta_1(t), \zeta_2(t), ..)^T$ is a white noise vector:

$$\langle \zeta_i(t), \zeta_j(\tau) \rangle = \delta_{ij}(t - \tau).$$

In *perturbation form*, the LNA is

$$\dot{x} = \widetilde{w}(x) + \varepsilon \tilde{\boldsymbol{G}}(x, \varepsilon) + \Omega^{-1/2} B(z)\Gamma, \tag{2.20}$$

with

$$x := \begin{pmatrix} z \\ X \end{pmatrix}, \quad \widetilde{w}(x) := \begin{pmatrix} N\mu \\ ND\mu \cdot X \end{pmatrix}, \quad \tilde{\boldsymbol{G}}(x, \varepsilon) := \begin{pmatrix} G(z, \varepsilon) \\ DG(z, \varepsilon) \cdot X \end{pmatrix}, \quad B := \begin{pmatrix} 0^{n \times m} \\ S\sqrt{F} \end{pmatrix}. \tag{2.21}$$

For a planar system in which $z \in \mathbb{R}^2$, the perturbation problem (2.20) has the form

$$\dot{x} = \boldsymbol{N}(x)\boldsymbol{\mu}(x) + \varepsilon \tilde{\boldsymbol{G}}(x, \varepsilon) + \Omega^{-1/2} B(z)\Gamma, \tag{2.22}$$

where $x = (x_1, x_2, X_1, X_2)^T$, and the critical set, $\widetilde{S}$, that emerges when $\varepsilon = \Omega^{-1} = 0$ is

$$\widetilde{S} := \{x \in \mathbb{R}^4 : \boldsymbol{\mu}(x) = 0\}. \tag{2.23}$$

The corresponding projection operator is

$$\widetilde{\Pi}^{\widetilde{S}_0} := \boldsymbol{I} - \boldsymbol{N}(\mathcal{D}\boldsymbol{\mu} \cdot \boldsymbol{N})^{-1} \mathcal{D}\boldsymbol{\mu}, \tag{2.24}$$

where $\mathcal{D}$ denotes differentiation with respect to x with $x := (x_1, x_2, X_1, X_2)^T$. The projection of the right hand side of (2.22) onto the tangent space of $\widetilde{S}_0$ is

$$\dot{x} = \varepsilon \widetilde{\Pi}^{\widetilde{S}_0} \tilde{\boldsymbol{G}}(x,0)|_{x\in\widetilde{S}_0} + \Omega^{-1/2} \widetilde{\Pi}^{\widetilde{S}_0} B(z) \Gamma|_{x\in\widetilde{S}_0}. \tag{2.25}$$

Remark 2.1. In the nonlinear Langevin regime, the reduced equation may contain a stochastic drift term that is $\mathcal{O}(\Omega^{-1})$ (see, Katzenberger [97] and Parsons and Rogers [132]). Hence, simply projecting onto the tangent space of the critical manifold does not yield a sufficient reduction of the Langevin equation unless the drift term vanishes or can be ignored. Such a term will also be present in the LNA regime. As Parsons and Rogers [132] point out, the drift term is not negligible when: the curvature of the slow manifold is significant, the curvature effect of the flow field is extreme, or the angle between the fast and slow subspace generates a bias in the way a trajectory returns to the slow manifold. It may be possible to discard the drift term when Ω is sufficiently large, but proof of this conjecture is open. Hence, (2.25) holds for systems that have a negligible (or identically zero) drift term.

2.3.4 The ssLNA for Systems in Standard Form: Comparison to Previous Results

The reduction method introduced by Thomas et al. [172] differs from (2.25). For a two-dimensional singularly perturbed problem in the standard form,[3]

$$\dot{x} = \varepsilon f(x,y), \tag{2.26a}$$

$$\dot{y} = g_0(x,y) + \varepsilon g_1(x,y,\varepsilon), \tag{2.26b}$$

the critical manifold $S_0 := \{(x,y) \in \mathbb{R}^2 : g_0(x,y) = 0\}$ attracts nearby trajectories if $g_{0y} < 0 \ \ \forall (x,y) \in S_0$. Moreover, by the Implicit Function Theorem, $g_{0y} \neq 0 \ \ \forall (x,y) \in S_0$

[3]For simplicity, we have assumed that $\varepsilon f(x,y)$ contains only terms that are $\mathcal{O}(\varepsilon)$, as this form is common in chemical kinetics. The analysis of the more general form can be found in Wechselberger [184].

implies the critical manifold is expressible as $y = h(x)$:

$$g_0(x, h(x)) = 0.$$

Thomas et al. [172] construct the ssLNA directly from the non-singular Jacobian[4] (that corresponds to $0 < \varepsilon$),

$$J = \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}, \quad \mathcal{S} = \begin{pmatrix} \mathcal{S}_{slow} \\ \mathcal{S}_{fast} \end{pmatrix}$$

with the a priori requirement that the system be in standard form. From this, they define the maps:

$$\bar{J} := f_x - f_y \cdot \frac{g_x}{g_y} \tag{2.27a}$$

$$\bar{\mathcal{S}}_{slow} := \mathcal{S}_{slow} - \frac{f_y}{g_y} \cdot \mathcal{S}_{fast} = \begin{pmatrix} 1 & -f_y/g_y \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{S}_{slow} \\ \mathcal{S}_{fast} \end{pmatrix} \tag{2.27b}$$

Let X and Y denote the respective fluctuations from the x and y. The ssLNA of Thomas et al. [170] is

$$\dot{x} = f(x, h(x)), \tag{2.28a}$$

$$\dot{X} = \bar{J}X + (\bar{\mathcal{S}}_{slow}\sqrt{F} \cdot \Gamma)|_{y=h(x)}. \tag{2.28b}$$

In contrast, to derive the deterministic sQSSA from GSPT, we begin with the *singular* Jacobian of the layer problem associated with (2.26). The corresponding projection operator Π^{S_0} is

$$\Pi^{S_0} := \begin{pmatrix} 1 & 0 \\ -g_{0x}/g_{0y} & 0 \end{pmatrix}, \tag{2.29}$$

and again the level set $g_0(x, y) = 0$ defines the critical manifold, S_0. The perturbation term,

[4]The notation f_y denotes $\partial_y f(x, y)$.

$\varepsilon G(z,\varepsilon)$ is

$$\varepsilon G(z,\varepsilon) := \varepsilon \begin{pmatrix} f(x,y) \\ g_1(x,y,\varepsilon) \end{pmatrix}$$

and therefore the reduced flow for the mean field is

$$\dot{x} = \begin{pmatrix} 1 & 0 \\ -g_{0x}/g_{0y} & 0 \end{pmatrix} \begin{pmatrix} f(x,y) \\ g_1(x,y) \end{pmatrix} = \begin{pmatrix} f(x,y) \\ -g_{0x}/g_{0y} \cdot f(x,y) \end{pmatrix}.$$

Again, $g_{0y} \neq 0 \ \ \forall z \in S_0$ implies $y = h(x)$ such that $g_0(x, h(x)) = 0$. Thus, the sQSSA for x is:

$$\dot{x} = f(x, h(x)). \tag{2.30}$$

For the corresponding ssLNA, and for two-dimensional systems of the standard form (2.26), we have

$$\boldsymbol{N}(x) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\mu}(x) = \begin{pmatrix} g_0(x,y) \\ g_{0x}X + g_{0y}Y \end{pmatrix},$$

and thus the critical manifold is

$$\widetilde{S} := \{(x,y,X,Y) \in \mathbb{R}^4 : g(x,y) = 0,\ g_{0x}X + g_{0y}Y = 0\}.$$

The perturbation term, $G(z,\varepsilon)$, is

$$G(z,\varepsilon) := \begin{pmatrix} f(x,y) \\ g_1(x,y,\varepsilon) \\ f_x(x,y)X + f_y(x,y)Y \\ g_{1x}(x,y,\varepsilon)X + g_{1y}(x,y,\varepsilon)Y \end{pmatrix}.$$

Computing $\widetilde{\Pi}^{\widetilde{S}_0}$ from $\boldsymbol{\mu}(x)$ and $\boldsymbol{N}(x)$ and projecting $G(z,0)$ onto the tangent space of the critical manifold yields

$$\dot{x} = f(x, h(x)) \tag{2.31a}$$

$$\dot{X} = \left(f_x - f_y \cdot \frac{g_{0x}}{g_{0y}}\right) X + (\Omega^{-1/2} \mathcal{S}_{slow} \sqrt{F} \cdot \Gamma)|_{y=h(x)}, \tag{2.31b}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\Big(f(x, h(x))\Big) X + (\Omega^{-1/2} \mathcal{S}_{slow} \sqrt{F} \cdot \Gamma)|_{y=h(x)}. \tag{2.31c}$$

Remark 2.2. We note that for systems in standard form, an equivalent reduction to (2.31) is given in Herath and Del Vecchio [74] and extends to non-autonomous systems.

Note the difference from the ssLNA of Thomas et al. [172]. First, we do not map

$$\mathcal{S}_{slow} \mapsto \mathcal{S}_{slow} - \frac{f_y}{g_y} \cdot \mathcal{S}_{fast}.$$

This is a consequence of the fact that our derivation from GSPT begins with the singular Jacobian, which is consistent with singular perturbation theory. In contrast, Thomas et al. [172] began with the perturbed, non-singular Jacobian. Consequently, when derived from GSPT, the ssLNA contains fewer diffusion terms than the ssLNA of Thomas et al. [172]. However, for planar systems $\mathcal{S}_{slow}$ and $\bar{\mathcal{S}}_{slow}$ should be close whenever $|f_y/g_y| \ll 1$. Hence, the difference between the ssLNA of Thomas et al. [172] and (2.31) should be small when the perturbation is in standard form. We note that $\mathcal{S}_{slow}$ is also invariant in the ssLNAs of Herath and Del Vecchio [74] and Pahlajani et al. [129].

Second, observe that

$$\bar{J} = f_x - f_y \cdot \frac{g_x}{g_y} \neq f_x - f_y \cdot \frac{g_{0x}}{g_{0y}}$$

unless $g(x,y)$ does not depend on ε, which is not always the case in applications. This difference follows from the utilization of the singular Jacobian in the derivation.

2.3.5 Benchmark Example: The MM Reaction Mechanism in the Limit of Small k_2 in (c,p) Coordinates

To demonstrate the projection operator methodology on a problem that is in standard form, we analyze the MM reaction mechanism in (p,c) coordinates and consider the limit of small k_2: $k_2 \mapsto \varepsilon\widehat{k}_2$. In (p,c)-coordinates, the deterministic rate equations are given by[5]

$$\begin{pmatrix} \dot{p} \\ \dot{c} \end{pmatrix} = N(z)\mu(z) + \varepsilon G(z) := \begin{pmatrix} 0 \\ 1 \end{pmatrix} (k_1(e_0 - c)(s_0 - c - p) - k_{-1}c) + \varepsilon \begin{pmatrix} \widehat{k}_2 c \\ -\widehat{k}_2 c \end{pmatrix}, \quad (2.32)$$

which is in the standard form (2.26); p is the slow variable and c is the fast variable. The projection matrix is[6]

$$\Pi^{S_0} := \begin{pmatrix} 1 & 0 \\ -(\mu_c)^{-1}\mu_p & 0 \end{pmatrix}, \quad (2.33)$$

and the critical manifold

$$S_0 := \{(p,c) \in \mathbb{R}^2 : \mu(p,c) = 0, \;\; 0 \le c \le e_0, \;\; 0 \le p \le s_0\}$$

is normally hyperbolic and attracting since

$$\langle D\mu, N\rangle = \mu_c = \frac{\partial}{\partial c}\left[k_1(e_0 - c)(s_0 - c - p) - k_{-1}c)\right] < 0, \quad \forall (c,p) \in S_0. \quad (2.34)$$

Since $\mu_c < 0 \;\; \forall (c,p) \in S_0$, it follows from the Implicit Function Theorem that the critical manifold is expressible as $c = y(p)$,

$$y(p) = \frac{k_2}{2}\left(s_0 + e_0 + K_S - p - \sqrt{(s_0 + e_0 + K_S - p)^2 - 4e_0(s_0 - p)}\right). \quad (2.35)$$

[5] In (2.32), s_0 denotes the total substrate.

[6] Again, μ_c denotes $\partial_c\mu(p,c)$ and μ_p denotes $\partial_p\mu(p,c)$.

The reduced equation for p is

$$\dot{p} = k_2 y(p). \tag{2.36}$$

We now turn to the reduction of the LNA. The complete perturbation form of the LNA is

$$\begin{pmatrix} \dot{p} \\ \dot{c} \\ \dot{X}_p \\ \dot{X}_c \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu(p,c) \\ D\mu(p,c)\cdot X \end{pmatrix} + \varepsilon\widehat{k}_2 \begin{pmatrix} c \\ -c \\ X_c \\ -X_c \end{pmatrix} + \Omega^{-1/2} B \cdot \Gamma, \tag{2.37}$$

where B is given by

$$B := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{\varepsilon\widehat{k}_2 c} \\ \sqrt{(k_1(e_0-c)(s_0-p-c)} & -\sqrt{k_{-1}c} & -\sqrt{\varepsilon\widehat{k}_2 c}, \end{pmatrix} \tag{2.38}$$

and $\mu(p,c)$ in (2.37) is

$$\mu(p,c) = k_1(e_0-c)(s_0-c-p) - k_{-1}c, \tag{2.39a}$$

$$D\mu(p,c)\cdot X = \mu_p X_p + \mu_c X_c. \tag{2.39b}$$

The critical manifold,

$$\widetilde{S} := \{(p,c,X_p,X_c) \in \mathbb{R}^4 : \mu(p,c) = 0 \ \ \& \ \ \mu_p X_p + \mu_c X_c = 0\}, \tag{2.40}$$

is normally hyperbolic and attracting. Proceeding in the usual way by calculating $\widetilde{\Pi}^{\widetilde{S}_0}$, the reduced equation for X_p is

$$\mathrm{d}X_p = k_2 X_c \mathrm{d}t + \Omega^{-1/2}\sqrt{k_2 y(p)}\ \mathrm{d}W_3(t). \tag{2.41}$$

To eliminate X_c from (2.41) we invoke the critical manifold relationship

$$X_c = -\mu_c^{-1}\mu_p X_p \;\text{ with }\; c = y(p), \tag{2.42}$$

which yields

$$\mathrm{d}X_p = k_2 y'(p) X_p \mathrm{d}t + \Omega^{-1/2}\sqrt{k_2 y(p)}\;\mathrm{d}W_3(t), \tag{2.43}$$

where "$y'(p)$" denotes $\dfrac{\mathrm{d}y}{\mathrm{d}p}$. Interestingly, it is worthwhile noting that equations (2.35) and (2.43) are equivalent to the tQSSA in the linear noise regime.

2.4 Reduction of the CME: Intimations from the Linear Noise Regime

In this section, we discuss the reduction of the CME for the MM reaction mechanism and its relationship to singular perturbations and critical manifolds. Specifically, we address the presence of transcritical bifurcations in the linear noise regime and illustrate that knowledge of the critical manifold can assist in avoiding erroneous conclusions concerning the validity of the stochastic QSSA.

2.4.1 Dynamic Bifurcations and the Segel–Slemrod sQSS Condition

The CME for the MM reaction mechanism is

$$\frac{\partial}{\partial t}P(n_S, n_C, t) = \Bigg[\frac{k_1}{\Omega}(\mathbb{E}_S^{+1}\mathbb{E}_C^{-1} n_S(n_{e_0} - n_C) + k_{-1}(\mathbb{E}_S^{+1}\mathbb{E}_C^{-1} - 1)n_C + k_2(\mathbb{E}_C^{-1} - 1)n_C\Bigg] P(n_S, n_C, t), \tag{2.44}$$

where n_{e_0} denotes the total number of enzyme molecules, and $P(n_S, n_C, t)$ is the probability of finding the system with n_S substrate molecules and n_C complex molecules at time t.

The homologous sQSS reduction of (2.44) is as follows. Given that there are n_S substrate molecules at time t, the probability that one one product molecule forms in an infinitesimal

window $[t, t + \mathrm{d}t)$ is

$$P(n_S - 1, t + \mathrm{d}t | n_S, t) := a(n_S)\mathrm{d}t = \frac{k_2 e_0 n_S}{K_M + n_S/\Omega}\,\mathrm{d}t, \tag{2.45}$$

where the propensity function, $a(n_S)$, is adopted from the deterministic sQSSA rate law, and the reduced CME is

$$\frac{\partial}{\partial t}P(n_S, t) = (\mathbb{E}_S^{+1} - 1)\frac{k_2 e_0 n_S}{K_M + n_S/\Omega}P(n_S, t). \tag{2.46}$$

In what follows, for simplicity, we set $\Omega = 1$ and work in arbitrary units (with a large number of molecules).

Numerical work by Sanft et al. [152] suggests that the Segel–Slemrod condition (expressed in terms of stochastic rate constants)

$$e_0 \ll K_M + s_0 \tag{2.47}$$

ensures the validity of the stochastic sQSSA (2.46). This is surprising, especially since the long-time validity of the deterministic sQSSA for the MM reaction mechanism requires $e_0/K_M \ll 1$ [42], which is more restrictive than the Segel and Slemrod condition. However, excellent (and extremely rigorous) work by Kang et al. [94] disputes this claim. In the stochastic regime, Kang et al. [94] concluded that $e_0 \ll K_M$, which is in agreement with the deterministic qualifier. A similar conclusion was drawn by Mastny et al. [117].

Importantly, although the deterministic and stochastic QSS reductions of the MM mechanism are justified via singular perturbation theory, the rigorous derivation of the sQSSA from singular perturbation was only recently established Goeke et al. [64]. This raises the question: given what we now understand about the bifurcation structure of the critical set associated with the deterministic MM reaction mechanism, what consequence(s) does this have on the stochastic QSS reduction? More specifically, does the Segel and Slemrod con-

dition guarantee that the stochastic sQSSA will remain accurate for all time, or is the more restrictive condition derived by Kang et al. [94] and Mastny et al. [117] necessary to ensure the accuracy of the stochastic sQSSA?

To answer this question, we note that the perturbation problem corresponding to small k_{-1} and k_2 is of the form (2.26):

$$\dot{p} = \varepsilon \widehat{k}_2 c \tag{2.48a}$$

$$\dot{c} = k_1(e_0 - c)(s_0 - c - p) - \varepsilon(\widehat{k}_{-1} + \widehat{k}_2). \tag{2.48b}$$

If $e_0 \ll s_0 = s(0)$, then the Fenichel reduction is formulated by projecting the perturbation onto the tangent space of $\mathcal{S}_a^r$, $T_z\mathcal{S}_a^r$:

$$\dot{p} = k_2 e_0, \tag{2.49a}$$

$$\dot{c} = 0. \tag{2.49b}$$

This approximation does not hold for all-time: eventually the trajectory follows $\mathcal{S}_a^b$, and the Fenichel reduction is

$$\dot{p} = k_2(s_0 - p), \tag{2.50a}$$

$$\dot{c} = -k_2(s_0 - p). \tag{2.50b}$$

The behavior of the reduction in small neighborhoods containing the bifurcation point is beyond the scope of this paper. In general, one must defer to non-classical methods to derive scaling laws near the bifurcation point (see, Berglund and Gentz [11], Krupa and Szmolyan [103]).

As Ω shrinks and fluctuations emerge, the LNA holds sway. The presence of a bifurcation point in the critical set is not too restrictive in this case. The ssLNA obtained via projection

onto $\mathcal{S}_a^r$ is

$$\dot{p} = k_2 e_0, \tag{2.51a}$$

$$\mathrm{d}X_p = \Omega^{-1/2}\sqrt{k_2 e_0}\,\mathrm{d}W_3. \tag{2.51b}$$

Note the relationship between the mean (2.51a) and variance (2.51b) as compared to (2.19). Likewise, projecting onto $T_z\mathcal{S}_a^b$ yields

$$\dot{p} = k_2(s_0 - p), \tag{2.52a}$$

$$\mathrm{d}X_p = -k_2 X_p \mathrm{d}t + \Omega^{-1/2}\sqrt{k_2(s_0 - p)}\,\mathrm{d}W_3. \tag{2.52b}$$

As $\Omega \to 0$ the CME prevails as the physically relevant model. The ssLNA (2.51) is a Gaussian process with equal mean and variance. In the CME regime, the reaction mechanism on $\mathcal{S}_a^r$ is equivalent to

$$\emptyset \xrightarrow{\lambda} \mathrm{P}, \tag{2.53}$$

where $\lambda = k_2 n_{e_0}$ and n_{e_0} denotes the total number (bound or unbound) of enzyme molecules. The CME that describes (2.53) is solvable. Let $P(N,t)$ denote the probability that there are N product molecules at time t. Then,

$$P(N,t) = \exp(-\lambda t) \cdot \frac{(\lambda t)^N}{N!}. \tag{2.54}$$

Note the consistency with (2.51). The Poisson jump process is approximately Gaussian when the system size is sufficiently large.

Unfortunately, (2.54) is not valid for all-time, and it is necessary to ascertain the range of its validity. More precisely, we ask: How long (on average) from the onset of the reaction does it take before (2.56) is valid? Since (2.51) is a Poisson process the jump times, t_N, are

gamma-distributed:

$$t_N \sim \lambda^N \exp(-\lambda t) \cdot \frac{t^{N-1}}{(N-1)!}.$$

Let n_{s_0} denote the total number of substrate molecules. The average time it takes to produce $N^* = n_{s_0} - n_{e_0}$ product molecules is

$$\langle t_{N^*} \rangle = \frac{N^*}{\lambda},$$

which is exactly homologous to the deterministic scenario.

Moving on, once $N = n_{s_0} - n_{e_0}$ we arrive at the intersection of the critical branches, $\mathcal{S}_a^r \cap \mathcal{S}_a^b$. At this point, no substrate molecules remain and the formation product is synonymous with the depletion of c:

$$C \xrightarrow{k_2} \mathrm{P}. \tag{2.55}$$

Once again, the CME associated with (2.55) is solvable:

$$P(n_{s_0} - n_c, t) = P(n_c, t) = \exp(-k_2 n_c t) \begin{pmatrix} n_{e_0} \\ n_c \end{pmatrix} (1 - \exp(-k_2 t))^{(n_{e_0} - n_c)}, \tag{2.56}$$

where n_c denotes the number of complex molecules, and time has been translated so that:

$$P(n_{s_0} - n_{e_0}, 0) = 1.$$

The question that remains is: How should the Gillespie algorithm be modified to reduce the computational complexity when k_2 and k_{-1} are sufficiently small? The Gillespie algorithm is employed to stochastically simulate chemical mechanisms in which it is assumed that the random time between two elementary reactions is exponentially distributed [62]. Here, these exponential distributions depend on the reaction rates. We recount the mathematical

Algorithm 2.1 Gillespie Algorithm [62]

Initialize system's state $Z(0) = Z_0$ at $t = 0$.
Compute propensity functions for each reaction, $a_j(Z(t))$ for system state $Z(t)$ at time t.
Compute sum of propensity functions

$$a_0(Z(t)) = \sum_{j=1}^{R} a_j(Z(t)).$$

Generate two random numbers r_1 and r_2 uniformly distributed in $(0, 1)$.
Compute

$$\tau = \frac{1}{a_0} \log\left(\frac{1}{r_1}\right)$$

such that the next chemical reaction occurs at time $t + \tau$.
Compute the smallest integer i such that

$$r_2 < \frac{\sum_{i'=1}^{i} a_{i'}(Z(t))}{a_0(Z(t))}.$$

$Z(t + \tau) \leftarrow Z(t) + \nu_i$
Repeat above steps until a stopping criterion is reached, for example, a maximum product concentration.

setup of the Gillespie Algorithm in Algorithm 2.1 where a_j is the propensity function and ν_j is the stoichiometric vector that correspond to the jth elementary reaction, see (2.8).

The above analysis indicates that $P(N, t)$ depends on the number of product molecules present at a given time in the reaction. Specifically, $P(N, t)$ depends on whether or not $N < N^*$.

To modify the Gillespie algorithm, observe that the propensity function $a(N)$ for product formation – at any given time – depends on the number of product molecules, N, present at time t. Thus,

$$a(N) = \begin{cases} k_2 n_{e_0}, & \text{if } N < N^*, \\ k_2 n_c, & \text{if } N \geq N^*. \end{cases} \tag{2.57}$$

Numerical simulations support the results of our analysis, and demonstrate that the Segel and Slemrod condition does not imply the validity of the stochastic sQSSA (see FIGURE 2.2).

We note that one can employ the reduction technique of Thomas et al. [172]. In general,

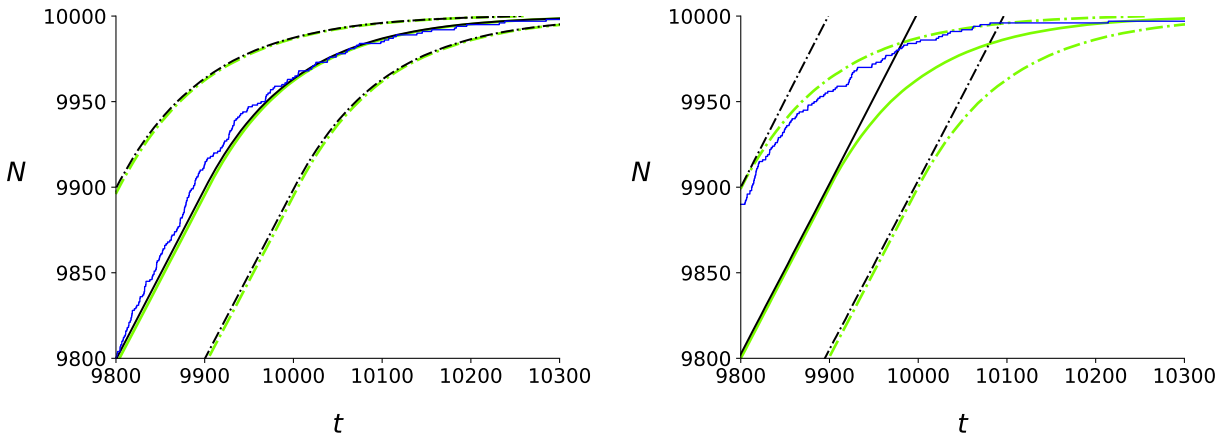


Figure 2.2: **The stochastic sQSSA (2.45) will fail near the bifurcation point if $e_0/K_M \gg 1$.** In both panels, the solid green curve is the numerically-computed mean of the timecourse for N (the number of product molecules) obtained from 1000 simulations generated by the Gillespie algorithm; the dashed/dotted green curves demarcate the mean $\pm$ the standard deviation. The solid black curve is the numerically-computed mean of the time course for N (the number of product molecules) obtained from 1000 simulations generated by the Gillespie algorithm equipped with a QSS-derived propensity function; the blue line shows one such randomly picked simulation; the dashed/dotted black lines demarcate the mean $\pm$ the standard deviation. In both simulations, the parameters (in arbitrary units) are: $n_{s_0} = 10000$, $n_{e_0} = 100$, $k_1 = 1000.0$, $k_{-1} = 0.01$, $k_2 = 0.01$ and $\Omega = 1.0$. $n_s(0) = n_{s_0}$, $n_c(0) = 0$, $n_e(0) = n_{e_0}$, and $N(0) = 0$. Note that $e_0/(K_M + s_0) \approx 0.01$ and therefore the Segel-Slemrod (2.47) condition holds. LEFT panel: The black solid and dashed/dotted lines are obtained from Gillespie algorithm equipped with the propensity function (2.57); the first two statistical moments are practically indistinguishable. RIGHT panel: The black solid and dashed/dotted lines are obtained from Gillespie algorithm equipped with the propensity function (2.45), and the stochastic sQSSA fails near $N = N^*$.

the ssLNA of Thomas et al. [172] will be close (in the asymptotic sense) to (2.51)–(2.52), but will be more complicated due to the presence of additional diffusion terms. The simplicity of the GSPT-derived ssLNA (2.51)–(2.52) helps to explain the insufficiency of the Segel–Slemrod condition for the validity of the stochastic sQSSA in the linear noise regime, thereby validating the results of Kang et al. [94] and Mastny et al. [117] from the context of GSPT.

As a final remark, we mention that the bifurcation point can also be handled with appropriate utilization of the tQSSA. Although the treatment of bifurcation points has so far not been addressed in the stochastic tQSSA literature, several rigorous studies suggest that the stochastic tQSSA is superior to the sQSSA in the CME and LNA regimes. Rigorous

analyses of the stochastic tQSSA are found in [99, 98, 100, 113, 7].

2.5 Discussion

The primary contribution of this work is the derivation of the ssLNA in a way that is consistent with Fenichel theory [55], that is, the projection of a perturbation term onto the tangent space of a normally hyperbolic critical manifold. Our derivation explains the origin of the differences between the the ssLNAs reported in Thomas et al. [172], Pahlajani et al. [129], and Herath and Del Vecchio [74].

By re-deriving the ssLNA directly from GSPT, we illustrated how the GSPT-derived ssLNA can be extended to singular perturbation problems where normal hyperbolicity fails and classical Fenichel theory breaks down. To the best of our knowledge, this is the first extension of the ssLNA to singular perturbation problems that contain a transcritical bifurcation.

Finally, let us remark on the possible special role of the standard form in the reduction of the CME. In their derivation of the ssLNA, Thomas et al. [172] shared the following insight: the mapping

$$\mathcal{S}_{slow} \mapsto \mathcal{S}_{slow} - \frac{f_y}{g_y} \cdot \mathcal{S}_{fast}$$

does not result in physically meaningful slow variable stoichiometry in the CME regime. However, as we pointed out, when the system is truly in standard form (again, the MM reaction mechanism with small e_0 does not technically qualify), the stoichiometry component of the slow variable, $\mathcal{S}_{slow}$ is invariant: $\mathcal{S}_{slow} \mapsto \mathcal{S}_{slow}$. This suggests that singularly perturbed systems in standard form[7] might, in some way, be amenable to QSS reduction in the CME regime. However, this hypothesis warrants further investigation.

[7] In addition, the drift term $\nu\mathcal{D}(x)$ introduced by Katzenberger [97] will presumably be zero when the system is in standard form.

CHAPTER 3

Accuracy versus Predominance: Reassessing the Validity of the Quasi-Steady-State Approximation

3.1 Introduction

The Michaelis–Menten (MM) reaction mechanism, a cornerstone of enzyme kinetics, describes the irreversible catalysis of a substrate, S, into product, P, by an enzyme, E:

$$\mathrm{S} + \mathrm{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_2} \mathrm{E} + \mathrm{P} \tag{3.1}$$

where C represents an enzyme-substrate intermediate complex [118]. The simplest form of the mass action equations modeling the dynamics of the substrate, complex, enzyme, and product concentrations (s, c, e, and p, respectively) is a two-dimensional system of ordinary differential equations (ODEs):

$$\dot{s} = -k_1(e_0 - c)s + k_{-1}c, \tag{3.2a}$$

$$\dot{c} = \quad k_1(e_0 - c)s - (k_{-1} + k_2)c, \tag{3.2b}$$

with $\dot{e} = -\dot{c}$, $\dot{p} = k_2 c$, where "$\dot{x}$" denotes differentiation of x with respect to time, t. The constant, e_0, is the total enzyme concentration, a conserved quantity.

Closed-form solutions to (3.2) are unattainable due to quadratic nonlinearities. Consequently, reduced equations approximating the long-time dynamics of (3.2) are often employed

to both elucidate the kinetics of (3.1) and serve as a forward model for parameter estimation from in vitro progress curve experiments via nonlinear regression [158, 168, 24]. The standard quasi-steady-state approximation (sQSSA) to the mass action equations associated with the MM reaction mechanism is a well-known and extensively studied reduced model in biochemical kinetics [158, 160]. Simply put, the sQSSA is a reduced system that approximates the depletion of substrate:

$$\dot{s} = -\frac{k_2 e_0 s}{K_M + s}, \quad s(0) = s_0, \tag{3.3}$$

where $K_M = (k_{-1} + k_2)/k_1$ is the Michaelis constant and $k_2\, e_0$ is the limiting rate. The Michaelis constant is composite of the equilibrium constant $K_S = k_{-1}/k_1$ and the Van Slyke-Cullen constant $K = k_2/k_1$, and can be expressed as $K_M = K_S + K$. We assume initial conditions for (3.2) lie on the s-axis, i.e., $(s, c)(0) = (s_0, 0)$. Under this assumption, we can equip (3.3) with the initial condition $s(0) = s_0$.[1]

Geometric singular perturbation theory provides mathematical justification for (3.3): with $(s, c)(0) = (s_0, 0)$, the solution to the first component of the mass action system (3.2a) converges to the approximation (3.3) in the limit as $e_0 \to 0$, provided all other parameters, k_{-1}, k_2 and k_1, are bounded above and below by positive constants [see, e.g. 65, 50]. For further information on geometric singular perturbation theory, see Kaper and Kaper [95], Hek [72], Jones [89] and Wechselberger [184].

While the mathematical rationale behind (3.3) is well-established, applying (3.3) presents a challenge: determining how small e_0 must be to ensure (3.3) is sufficiently accurate. This assessment is complicated by the relative nature of "small." e_0 is only small compared to another quantity, but what is that quantity? Emerging research [see, e.g. 49, 51, 50] suggests that the sQSSA is accurate whenever $e_0 \ll K_M$, or $\varepsilon_{RS} = e_0/K_M \ll 1$, which we refer to as the Reich-Sel'kov condition [145]. Thus, the "other" quantity is the Michaelis constant,

[1] Loosely speaking, solutions to the mass action system converge to solutions of the reduced system as $\varepsilon \to 0$ as long as initial conditions are sufficiently close to the critical manifold. See A.2 for details on the initial conditions of quasi-steady-state reductions.

K_M, and the sQSSA is considered *valid* – in the sense of accuracy – when the Reich-Sel'kov condition holds.

Beyond accuracy, it is crucial to determine conditions that ensure the *predominance* of the sQSSA (3.3). When the reaction unfolds away from the singular limit (i.e., with $0 < e_0$), multiple reduced models may be accurate. However, it is not obvious that the Reich-Sel'kov condition guarantees that the sQSSA is the predominant reduction among the nearby Fenichel reductions. Other reduced models may be more accurate[2] than the sQSSA even when $e_0 \ll K_M$. Therefore, identifying the most accurate approximation among the available reduced models with overlapping validity conditions is essential, and this most accurate model should be labeled as the predominant reduction (see Section 3.2 for a better understanding of the class of reduced models considered).

This chapter aims to deepen our understanding of the Reich-Sel'kov condition and distinguish between accuracy and predominance concerning the sQSSA (3.3). To our knowledge, this distinction has not been previously addressed in the literature. Using rigorous phase-plane analysis and methods from ODE analysis, we demonstrate that while the Reich-Sel'kov condition provides a general indication of the sQSSA's accuracy, it does not ensure its predominance, which requires a more restrictive criterion. We derive this qualifier and challenge the notion that the validity of the sQSSA solely pertains to its accuracy. We argue that validity should encompass both the accuracy and predominance of the sQSSA. In essence, the qualifier establishing the sQSSA's validity is more restrictive than any previously reported condition.

3.2 The Open Problem: Identifying the Predominant Reduction

Fenichel theory provides quasi-steady-state reductions for certain TFPVs in the parameter space. To reiterate, for the MM system (3.2), the parameter space is a subset of $\mathbb{R}^4_{\geq 0}$ with $p = (e_0, k_1, k_{-1}, k_2)^T$. Here, $p^* := (0, k_1, k_{-1}, k_2)^T$ is a TFPV if k_1, k_{-1}, k_2 are bounded by

[2]In this context, *accuracy* refers to the ability of the reduction to approximate the flow on the slow manifold. We are choosing to ignore errors that arise in the approach to the slow manifold.

positive constants. When $p = p^*$, the critical manifold, S_0, is the s-axis, which is normally hyperbolic and attractive:

$$S_0 := \{(s, c) \in \mathbb{R}^2_{\geq 0} : c = 0\}. \tag{3.4}$$

Thus, Fenichel theory yields the standard quasi-steady-state approximation for small e_0:

$$\begin{pmatrix} \dot{s} \\ \dot{c} \end{pmatrix} = -\frac{\varepsilon \widehat{e}_0 k_2 s}{K_M + s} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv -\frac{k_2 e_0 s}{K_M + s} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{3.5}$$

Other TFPV values exist for (3.2). For instance, with $p^* = (e_0, k_1, k_{-1}, 0)^T$, the critical manifold becomes

$$S_0 := \left\{ (s, c) \in \mathbb{R}^2_{\geq 0} : c = \frac{e_0 s}{K_S + s} \right\} \tag{3.6}$$

and the corresponding Fenichel reduction is the slow product QSS reduction:

$$\begin{pmatrix} \dot{s} \\ \dot{c} \end{pmatrix} = -\frac{k_2 e_0 s (K_S + s)}{K_S e_0 + (K_S + s)^2} \begin{pmatrix} 1 \\ \dfrac{K_S e_0 - K(K_S + s)}{(K_S + s)^2} \end{pmatrix}. \tag{3.7}$$

This reduction corresponds to slow accumulation of product due to small k_2 and differs from (3.5). We will be analyze this reduction in Section 3.4.

Just as important as what Fenichel and TFPV theory tell us is what they do *not*. While the theory indicates that (3.5) approximates (3.2) as $e_0 \to 0$, it does not specify how small e_0 must be for (3.5) to be valid in practice, where the singular limit is never truly reached. In other words, how small is small enough?

The size of e_0 must be considered relative to other parameters. Fenichel theory requires eigenvalue disparity near the critical manifold. For small e_0, the ratio of slow and fast

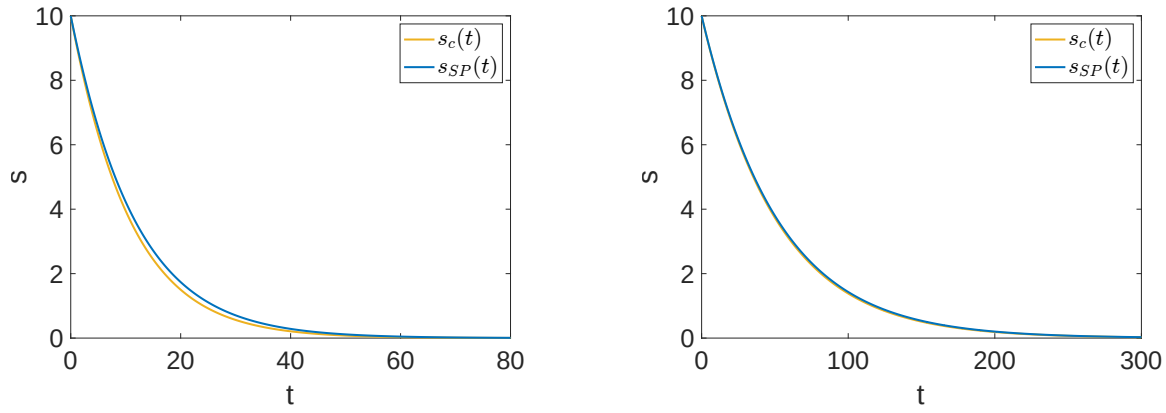


Figure 3.1: **The sQSSA** (3.5) **and the slow product QSSA** (3.7) **are indistinguishable for large** k_{-1}**. Left**: $s_0 = 10$, $c_0 = 0$, $e_0 = 10$, $k_1 = 1$, $k_{-1} = 100$, $k_2 = 1$. **Right**: $s_0 = 10$, $c_0 = 0$, $e_0 = 10$, $k_1 = 1$, $k_{-1} = 500$, $k_2 = 1$. Substrate depletion over time is shown in both panels. $s_c(t)$ is the numerical solution to (3.5) and $s_{SP}(t)$ is the numerical solution to (3.7).

eigenvalues near $(s, c) = (0, 0)$ is:

$$\delta := \frac{e_0}{K_M} \cdot \frac{k_2}{k_{-1} + k_2} + \mathcal{O}(e_0^2). \tag{3.8}$$

While $\delta \ll 1$ is necessary, it is not sufficient for the accuracy of (3.5) [49]. For instance, consider the case where $e_0 = K_M$ but $k_2 \ll k_{-1}$. This scenario renders small δ, but the tangent vector to the graph of the c-nullcline at the origin will not align with the slow eigenvector, which is crucial for long-time accuracy. This misalignment can lead to significant errors in the sQSSA approximation, even when δ appears to indicate otherwise.

The more restrictive condition $e_0 \ll K_M$, introduced in [145], has been rigorously proven sufficient to guarantee the accuracy of the sQSSA [42, 51, 94, 117, 48]. However, this condition does not fully address the issue of predominance. For instance, if k_{-1} is large and e_0 is reduced such that $k_1 e_0 \ll k_{-1}$ (implying $e_0 \ll K_M$), convergence to (3.5) is not implied. In fact, as $k_{-1} \to \infty$, the resulting singular perturbation problem yields a trivial QSSA: $\dot{s} = \dot{c} = 0$ (see A.1 for the formal calculation).

This example demonstrates that different TFPVs can share a critical manifold and be close in parameter space. For instance, consider a scenario where $k_1 e_0$ and k_2 are relatively

small compared to k_{-1}. In this case, the TFPVs $(0, k_1, k_{-1}, k_2)^T$ or $(e_0, k_1, k_{-1}, 0)^T$ could be quite close to each other. Terminating the approach to one TFPV might inadvertently lead to a point in parameter space that is actually closer to another TFPV.

Numerical simulations (FIGURE 3.1) illustrate this phenomenon. When k_{-1} is significantly large, the trajectories of the system under the sQSSA (3.5) and the slow product QSSA (3.7) become virtually indistinguishable. This makes it challenging to definitively determine which reduction is more accurate or appropriate for the given parameter values.

Moreover, the presence of multiple TFPVs introduces ambiguity in assessing the validity of the sQSSA. For example, if $e_0 = 1$, $k_1 = 0.1$, $k_{-1} = 10$, and $k_2 = 0.1$, it is unclear whether the system's behavior is better approximated by the sQSSA associated with the TFPV $(0, k_1, k_{-1}, k_2)^T$ or the slow product QSSA associated with $(e_0, k_1, k_{-1}, 0)^T$. This ambiguity arises because the parameter values do not fall neatly within the domain of a single TFPV, making it difficult to determine which reduction provides the most accurate representation of the system dynamics. Thus, we define the *predominance* of a given QSS reduction among the class of QSS approximations, obtained through Fenichel theory, for the nearby TFPVs. Well-known validity conditions such as the Reich-Sel'kov condition can typically guide us to find TFPVs with overlapping validity conditions for a robust analysis of predominance. However, more generally, determining proximity to different TFPVs is challenging due to the different units of the parameters. A dimensionless indicator is needed.

Remark 3.1. The approximation of the slow manifold is a well-studied field, with methods extending beyond Fenichel's reduction theory. The Fraser iterative method, for instance, provides approximations for slow manifolds with small curvatures [60, 124, 95]. In such cases, incorporating additional iterative terms can improve the accuracy of the approximations. Furthermore, several studies employ integration-based methods to generate numerical approximations of the slow flow within a specified error tolerance [29]. In our work, we adhere to other Fenichel reductions to examine the predominance of the standard QSSA, as their convergence to the slow flow has been extensively established for the Michaelis–Menten

mechanism.

The fact that the magnitude of e_0/K_M alone does not ensure the sQSSA (3.5) is the *predominant* reduction is a consequence of having to always operate away from the singular limit in a large parameter space. This naturally leads to the question: does there exist a dimensionless parameter whose magnitude not only ensures the accuracy of the sQSSA (3.3) but also its predominance among other known QSSAs? In other words, can we define a more comprehensive and robust notion of "validity" that encompasses both accuracy and predominance?

3.3 Anti-Funnels and the Slow Invariant Manifold

As discussed in Section 3.2, a central challenge in applying Fenichel theory to the MM system is the potential ambiguity arising from different TFPVs sharing the same critical manifold. This can lead to uncertainty about the validity of a specific reduction away from the singular limit. To address this, we turn to qualitative methods that allow us to determine the location of the slow manifold, S_0^ε, relative to known curves in the phase-plane. This section introduces the anti-funnel theorem, adapted from Hubbard and West [80] and following the approach of Calder and Siegel [22]. We outline a strategy to determine when the sQSSA (3.3) is not only accurate but also predominant.

Definition 3.1. Fences and anti-funnels. Consider the first-order differential equation $y' = f(x, y)$ over the interval $x \in I = [a, b)$ where $a < b \leq \infty$ and " $'$" denotes differentiation with respect to x. Let α and β be continuously-differentiable functions that satisfy

$$\alpha'(x) \leq f(x, \alpha(x)) \text{ and } f(x, \beta(x)) \leq \beta'(x) \tag{3.9}$$

for all $x \in I$.

(a) The curve α is a *lower fence* and the curve β is an *upper fence*.

(b) The lower and upper fences are *strong fences* if the respective inequalities are always

strict.

(c) The set

$$\Gamma := \{(x, y) \ : \ x \in I, \ \beta(x) \leq y \leq \alpha(x)\} \tag{3.10}$$

is called an *anti-funnel*[3] if

$$\beta(x) < \alpha(x) \ \text{ for all } \ x \in I.$$

(d) The anti-funnel is *narrowing* if

$$\lim_{x \to b^-} |\alpha(x) - \beta(x)| = 0. \tag{3.11}$$

Theorem 3.1. Anti-funnel Theorem [80]. Consider the first-order differential equation $y' = f(x, y)$ over the interval $I = [a, b)$ where $a < b \leq \infty$. If Γ is an *anti-funnel* with a strong lower fence, α, and a strong upper fence, β, then there exists a solution $y(x)$ to the differential equation such that

$$\beta(x) < y(x) < \alpha(x) \ \text{ for all } \ x \in I. \tag{3.12}$$

Furthermore, if Γ is *narrowing* and $\frac{\partial f}{\partial y}(x, y) \geq 0$ in Γ, then the solution $y(x)$ is unique.

In general, an anti-funnel is the region above an upper fence and below a lower fence in a two-dimensional phase space. Typically, anti-funnels (or funnels) are visualized with the independent variable increasing from left to right with the right endpoint of interval I being the end of the anti-funnel.

However, for the Michaelis–Menten system, the substrate concentration, s, decreases over time, with $s(0) = s_0$ and $s(t) \to 0$ as $t \to \infty$. In the (s, c) phase-plane, solution trajectories

[3]In the (x, y) phase plane, if the upper fence lies above the lower fence, the set between the two fences is known as a *funnel*. Conversely, if the lower fence lies above the upper fence, the set between the two fences is called an *antifunnel* [80]. See Hubbard and West [80] for the detailed theory.

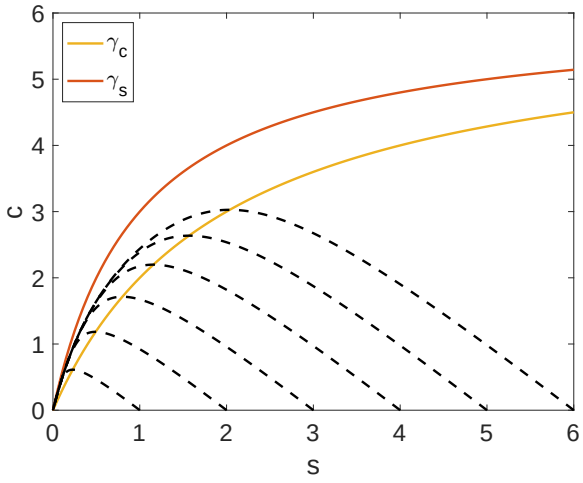


Figure 3.2: **The region, Γ, between the horizontal nullcline and the vertical nullcline curves is a trapping region on the phase plane.** The numerical solutions to the mass action equations (3.1) for several initial conditions are denoted by black dashed lines. $s_0 \in \{1, 2, 3, 4, 5, 6\}$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 1$, $k_2 = 1$. The numerical solution of trajectories start on the s-axis and eventually enter Γ.

start at the substrate axis and move from right to left with respect to the s-axis. Thus, the vector field always has a negative component in the s-direction.

Despite this difference in directionality, the concept of anti-funnels remains applicable. The vector field can still point inwards at the anti-funnel fences in the (s, c) coordinate system as we move from right to left, while satisfying the conditions in Definition 3.1. Consequently, solution trajectories can enter the anti-funnels defined in this chapter.

For the MM equations (3.2), we have:

$$\frac{\mathrm{d}c}{\mathrm{d}s} = -\frac{e_0 s - (K_M + s)c}{e_0 s - (K_S + s)c} =: f(c, s). \tag{3.13}$$

The distinguished solution to (3.13), "$c = y(s)$", satisfying (3.12) represents the unique slow manifold, S_0^ε. In context of an anti-funnel, the slow manifold would lie above an upper fence and below a lower fence in the phase plane. The challenge lies in finding suitable lower and upper fences, $\alpha(s)$ and $\beta(s)$.

A natural choice for the upper fence is the c-nullcline, γ_c, given by:

$$\text{Graph}(\gamma_c) = \{(s, c) \in \mathbb{R}^2 : c = e_0 s/(K_M + s)\} \tag{3.14}$$

Considering the biochemically relevant portion of $\mathbb{R}^2$, we focus on $\text{Graph}(\gamma_c) \cap \mathbb{R}^2_{\geq 0}$. This choice is advantageous because it represents the quasi-steady-state variety corresponding to the standard reduction (3.3), and all phase-plane trajectories starting on the s-axis eventually cross it (see Calder and Siegel [22] and Noethen and Walcher [125] for a proof of this statement).

The set Γ enclosed between the c-nullcline and the s-nullcline, γ_s,[4] is positively invariant[5]:

$$\Gamma := \{(s, c) \in \mathbb{R}^2_{\geq 0} \ : \ \gamma_c(s) \leq c \leq \gamma_s(s)\} \tag{3.15}$$

Once a trajectory enters Γ, it remains within Γ as $s \to 0$ or $t \to \infty$. (see Calder and Siegel [22] and Noethen and Walcher [125] for a detailed proof). FIGURE 3.2 illustrates this behavior.

Constructing a suitable lower fence requires more analysis. The stationary point $(s, c) = (0, 0)$ is a node with eigenvalues $\lambda_- < \lambda_+ < 0$:

$$\lambda_\pm = \frac{k_1}{2}(K_M + e_0)\left(-1 \pm \sqrt{1 - \frac{4Ke_0}{(K_M + e_0)^2}}\right). \tag{3.16}$$

Trajectories eventually approach the stationary point along the slow eigenvector, v^+, spanning the one-dimensional subspace:

$$T_0 S_0^\varepsilon = \text{span}(v^+) = \{(s, c) \in \mathbb{R}^2 : c = ms\}, \tag{3.17}$$

[4]The graph of the s-nullcline is the set $\{(s, c) \in \mathbb{R}^2 : c = e_0 s/(K_S + s)\}$,

[5]In this chapter, we refer to 'positive invariance' of sets in the phase-plane with respect to time. As $t \to \infty$, solutions that enter positive invariant sets will remain in them. We choose this terminology because time defines the progression of chemical reactions here.

with slope m:

$$m = \frac{1}{2k_{-1}}\left(-k_{-1} - k_2 + k_1 e_0 + \sqrt{(k_{-1} + k_2 + k_1 e_0)^2 - 4k_1 k_2 e_0}\right). \tag{3.18}$$

Calder and Siegel [22] leverage this property and define the lower fence:

$$\alpha_{cs}(s) := \frac{m e_0 s}{e_0 + ms}, \tag{3.19}$$

and prove that S_0^ε lies between $\gamma_c(s)$ and $\alpha_{cs}(s)$. Notably, the graph of α_{cs} lies below that of γ_s. By proving that

$$\Gamma_{cs} := \{(s, c) \ : \ s \in I, \gamma_c(s) \leq c \leq \alpha_{cs}(s)\} \tag{3.20}$$

is a narrowing anti-funnel,[6] Calder and Siegel [22] prove that the distinguished slow manifold, S_0^ε, also lies in Γ, the region contained between the c- and s-nullcline.

A key feature of (3.19) is $\alpha'_{cs}(0) = m$. This leads to highly accurate reduced equation for $\dot{s}$ near the stationary point:[7]

$$\dot{s} = -\frac{e_0 s}{e_0 + ms} \cdot (k_1 e_0 - k_{-1} m). \tag{3.21}$$

Despite its advantages, this approach has limitations. First, the reduction (3.21) may be unreliable away from the stationary point, even when $\lambda_- \ll \lambda_+ < 0$ (see Eilertsen et al. [47]). Second, the anti-funnel construction leads to the following:

Proposition 3.1. For (3.2) with initial condition $(s, c)(0) = (s_0, 0)$, let $t = t_{\text{cross}}$ be the time the trajectory crosses the upper fence $c = \gamma_c(s)$. Then, the bound

$$-\frac{k_2 e_0 s}{K_M + s} \leq \dot{s} \leq -\frac{e_0 s}{e_0 + ms} \cdot (k_1 e_0 - k_{-1} m) \tag{3.22}$$

[6] Calder and Siegel [22] define I as $[a, \infty)$ where $a > 0$ is arbitrary.

[7] The equation is obtained via direct substitution of $c = \alpha_{cs}(s)$ into (3.2a).

holds for all $t \geq t_{\text{cross}}$.

Proof. Since Γ_{cs} is a narrowing anti-funnel, S_0^ε lies between the graphs of $\alpha_{cs}(s)$ and $\gamma_c(s)$. As S_0^ε is invariant, any trajectory entering Γ_{cs} cannot leave it. Therefore, $\beta(s) \leq c \leq \alpha_{cs}(s)$ holds for all $t \geq t_{\text{cross}}$ since Γ is positively invariant with respect to time and $\Gamma_{cs} \subset \Gamma$. With $\dot{s} = -k_1 e_0 s + k_1(K_S + s)c$, it follows that

$$-k_1 e_0 s + k_1(K_S + s)\gamma_c(s) \leq \dot{s} \leq -k_1 e_0 s + k_1(K_S + s)\alpha_{cs}(s)$$

proving the assertion. □

While the upper bound in (3.22) is sharper than $\dot{s} \leq 0$, extracting quantitative information about the sQSSA's accuracy and dimensionless parameters like $\varepsilon_{RS} = e_0/K_M$ from (3.22) is not straightforward.

The goal is to define an anti-funnel with fences that provide both qualitative and quantitative information about S_0^ε and the error for a given QSSA. By "trapping" S_0^ε between suitable upper and lower fences that form narrowing anti-funnels, we can derive qualifiers that determine the accuracy and predominance of various QSSAs.

3.4 Trapping the Slow Manifold: The Standard QSSA

As stated in Section 3.3, our strategy is to construct upper and lower fences that form a narrowing anti-funnel. This raises the question: How do we construct suitable fences? One approach is approximating the slow manifold, S_0^ε, via perturbation expansion

$$S_0^\varepsilon = h_0(s) + \varepsilon h_1(s) + \varepsilon^2 h_2(s) + \mathcal{O}(\varepsilon^2). \tag{3.23}$$

Insertion of (3.23) into the invariance equation results in a *regular* perturbation problem. However, the coefficients in (3.23) depend on ε, which in turn depends on a specific TFPV. Since we aim for general error bounds, we need fences that are not heavily ε-dependent.

An alternative approach involves combining quantitative reasoning with creative insights. Kumar and Josić [108] sought to approximate the flow on S_0^ε for small k_2 ($k_2 \mapsto \varepsilon\widehat{k_2}$). They reasoned that

$$c = \frac{e_0 s}{K_M + s} \tag{3.24}$$

is a good approximation to S_0^ε when k_2 is small. Under this assumption, the flow of c on S_0^ε is determined by:

$$\dot{c} \approx \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{e_0 s}{K_M + s}\right)\dot{s} = \frac{K_M e_0}{(K_M + s)^2} \cdot \dot{s}. \tag{3.25}$$

Substituting (3.25) and (3.24) into the conservation law, $\dot{s} + \dot{c} + k_2 c = 0$, and solving for $\dot{s}$ yields:

$$\dot{s} = -\frac{k_2 e_0 s(K_M + s)}{K_M e_0 + (K_M + s)^2}. \tag{3.26}$$

This reduction is significant for two reasons. First, it effectively captures the Fenichel reduction for this scenario (compare to (3.7) with small k_2). Second, the graph of $\gamma(s)$, where

$$\gamma : s \mapsto \frac{e_0 s}{K_S + s} - \frac{k_2 e_0 s(K_M + s)}{k_1(K_S + s)(K_M e_0 + (K_M + s)^2)}, \tag{3.27}$$

is an excellent candidate for an appropriate anti-funnel fence.

As shown in Figure 3.3, γ_c, γ_s and γ are strictly increasing, with $\gamma_c < \gamma < \gamma_s$ for all $s > 0$. The set contained between the graphs of γ_c and γ is positively invariant:

$$\Gamma_0 := \{(s, c) \in \mathbb{R}^2_{\geq 0} \ : \ \gamma_c(s) \leq c \leq \gamma(s)\}. \tag{3.28}$$

Let t_{cross} be the time a trajectory crosses the c-nullcline, as established in [125, 22]. Then, $(s, c)(t) \in \Gamma_0$ for all $t \geq t_{\text{cross}}$. Importantly:

Theorem 3.2. For the differential equation (3.13), Γ_0 is a positively invariant, narrowing anti-funnel containing a unique slow manifold, S_0^ε.

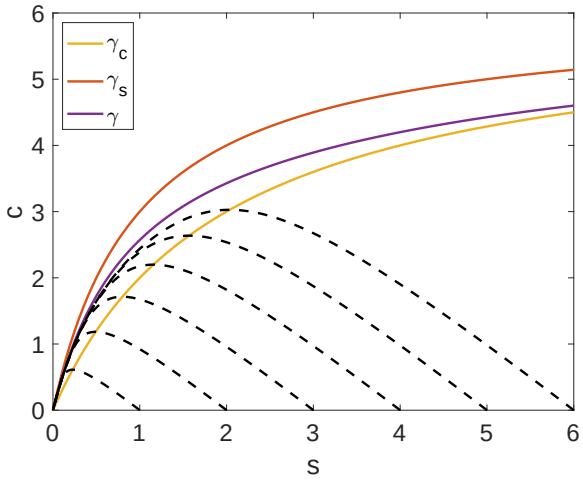


Figure 3.3: **The region between the horizontal nullcline curve $\gamma_c(s)$ and the curve $\gamma(s)$ is a trapping region on the phase plane.** The numerical solutions to the mass action equations (3.1) for several initial conditions are denoted by black dashed lines. $s_0 \in \{1, 2, 3, 4, 5, 6\}$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 1$, $k_2 = 1$.

Proof. To prove Γ_0 is positively invariant, we show that the vector field for (3.2) points inwards at the boundary curves:

$$\dot{c}(s, \gamma(s)) \ - \gamma'(s)\dot{s}(s, \gamma(s)) \ < 0, \tag{3.29a}$$

$$\dot{c}(s, \gamma_c(s)) - \gamma_c'(s)\dot{s}(s, \gamma_c(s)) > 0. \tag{3.29b}$$

We begin by evaluating the dynamics at the curve $c = \gamma(s)$:

$$\dot{s}(s, \gamma(s)) = -k_1 e_0 s + k_1(K_S + s)\gamma(s), \tag{3.30a}$$

$$= -k_1 e_0 s + k_1(K_S + s)\left(\frac{e_0 s}{K_S + s} - \frac{k_2 e_0 s(K_M + s)}{k_1(K_S + s)(K_M e_0 + (K_M + s)^2)}\right), \tag{3.30b}$$

$$= -\frac{k_2 e_0 s(K_M + s)}{K_M e_0 + (K_M + s)^2}, \tag{3.30c}$$

and

$$\dot{c}(s,\gamma(s)) = k_1 e_0 s - k_1(K_M+s)\gamma(s), \tag{3.31a}$$

$$= k_1 e_0 s - k_1(K_M+s)\left(\frac{e_0 s}{K_S+s} - \frac{k_2 e_0 s(K_M+s)}{k_1(K_S+s)(K_M e_0+(K_M+s)^2)}\right), \tag{3.31b}$$

$$= -\frac{k_2 e_0 s}{K_S+s} + \frac{k_2 e_0 s(K_M+s)^2}{(K_S+s)(K_M e_0+(K_M+s)^2)}. \tag{3.31c}$$

The derivative of $\gamma(s)$ is:

$$\gamma'(s) = \frac{K_S e_0}{(K_S+s)^2} - \frac{k_2 K_S e_0(K_M+s)}{k_1(K_S+s)^2(K_M e_0+(K_M+s)^2)} \tag{3.32}$$

$$- \frac{k_2 e_0 s}{k_1(K_S+s)(K_M e_0+(K_M+s)^2)} + \frac{2k_2 e_0 s(K_M+s)^2}{k_1(K_S+s)(K_M e_0+(K_M+s)^2)^2}. \tag{3.33}$$

The required expression at $c=\gamma(s)$ is:

$$\dot{c} - \gamma'(s)\dot{s} = -\frac{k_2 e_0 s}{K_S+s} + \frac{k_2 e_0 s(K_M+s)^2}{(K_S+s)(K_M e_0+(K_M+s)^2)} + \frac{k_2 e_0 s(K_M+s)}{K_M e_0+(K_M+s)^2}\gamma'(s), \tag{3.34a}$$

$$= -\frac{k_2 e_0 s}{(K_S+s)(K_M e_0+(K_M+s)^2)}\left(K_M e_0 - (K_M+s)(K_S+s)\gamma'(s)\right). \tag{3.34b}$$

To prove inequality (3.29a), we need to show that $\delta(s) = K_M e_0 - (K_M+s)(K_S+s)\gamma'(s) > 0$. This also proves that $\gamma(s)$ is a lower fence. Substituting $\gamma'(s)$ and simplifying, we obtain:

$$\begin{aligned}
\delta(s) &= K_M e_0 - \frac{K_S e_0(K_M+s)}{(K_S+s)} + \frac{k_2 K_S e_0(K_M+s)^2}{k_1(K_S+s)(K_M e_0+(K_M+s)^2)} \\
&\quad + \frac{k_2 e_0 s(K_M+s)}{k_1(K_M e_0+(K_M+s)^2)} - \frac{2k_2 e_0 s(K_M+s)^3}{k_1(K_M e_0+(K_M+s)^2)^2}, \\
&= \frac{k_2 e_0}{k_1(K_S+s)(K_M e_0+(K_M+s)^2)^2}\Big((K_M e_0+(K_M+s)^2 s)^2 \\
&\quad + K_S(K_M+s)^2(K_M e_0+(K_M+s)^2) \\
&\quad + (K_M+s)(K_S+s)(K_M e_0+(K_M+s)^2)s - 2(K_M+s)^3(K_S+s)s\Big).
\end{aligned}$$

Expanding the first term in the expression,

$$\begin{aligned}\delta(s) &= \frac{k_2 e_0}{k_1(K_S+s)(K_M e_0+(K_M+s)^2)^2}\Big((K_M e_0)^2 s + 2(K_M e_0)(K_M+s)^2 s + (K_M+s)^4 s \\ &\quad + K_S(K_M+s)^2(K_M e_0+(K_M+s)^2) \\ &\quad + (K_M+s)(K_S+s)(K_M e_0+(K_M+s)^2)s - 2(K_M+s)^3(K_S+s)s\Big), \\ &= \frac{k_2 e_0}{k_1(K_S+s)(K_M e_0+(K_M+s)^2)^2}\Big((K_M e_0)^2 s + 2(K_M e_0)(K_M+s)^2 s \\ &\quad + (k_2/k_1)(K_M+s)^3 s + K_S(K_M+s)^2(K_M e_0+(K_M+s)^2) \\ &\quad + K_M e_0(K_M+s)(K_S+s)s\Big),\end{aligned}$$

This expression is a sum of positive terms, proving inequality (3.29a).

For inequality (3.29b), a straightforward calculation at $c = \gamma_c(s)$ gives:

$$\dot{c} - \gamma_c'(s)\dot{s} = 0 + \left(\frac{K_M e_0}{(K_M+s)^2}\right)\left(\frac{k_2 e_0 s}{K_M+s}\right)$$

which is always positive for $s > 0$. Thus, Γ_0 is a positively invariant region.

Positive invariance implies trajectories entering Γ_0 remain within. Calder and Siegel [22], as well as Noethen and Walcher [125], prove that for solutions initializing on the s-axis, there exists a $t_{\text{cross}} > 0$ at which they enter Γ_0 by crossing γ_c.

From (3.29), γ and γ_c are strong lower and upper fences, respectively, with $\gamma(s) > \gamma_c(s)$ for $s > 0$. Furthermore,

$$\lim_{s\to\infty} |\gamma(s) - \gamma_c(s)| = 0 \quad \text{and} \quad \frac{\partial f}{\partial c}(c,s) \geq 0 \tag{3.35}$$

in Γ_0. The assertion follows from Theorem 3.1. □

Theorem 3.2 establishes that Γ_0 contains the slow manifold S_0^ε, independent of any specific perturbation scenario. This allows us to extract qualitative and quantitative information about the accuracy of a QSSA and assess predominance when multiple TFPVs are close.

The rest of this section is organized as follows. In Section 3.4.1, we use the improved trapping region to derive a quantitative error estimate for the sQSS approximation and recover the Reich-Sel'kov condition. In Section 3.4.2, we explore the conditions where both the sQSS and reverse QSS reductions are valid, examining their validity as we move away from the Reich-Sel'kov condition.

3.4.1 Accuracy of the Standard QSSA

To leverage Γ_0's properties, we introduce sharp bounds on the substrate depletion for $t \geq t_{\text{cross}}$, representing the slow regime.

Proposition 3.2. For system (3.2) with initial condition $(s, c)(0) = (s_0, 0)$, the bound

$$-\frac{k_2 e_0 s}{K_M + s} \leq \dot{s} \leq -\frac{k_2 e_0 s(K_M + s)}{K_M e_0 + (K_M + s)^2} \tag{3.36}$$

holds for all $t \geq t_{\text{cross}}$.

Proof. The proof follows from the construction of Γ_0 and substituting $\gamma_c(s) \leq c \leq \gamma(s)$ into (3.2a). □

Since the LHS of (3.36) is the sQSSA for substrate depletion, the design of Γ_0 suggests that the sQSSA performs well when γ and γ_c are close. Rewriting (3.36) as:

$$-\frac{k_2 e_0 s}{K_M + s} \leq \dot{s} \leq -\frac{k_2 e_0 s}{K_M + s}\left(\frac{1}{1 + \varphi(s)}\right), \quad \varphi(s) = \frac{K_M e_0}{(K_M + s)^2} \tag{3.37}$$

we see that, for any subinterval $\ell \subset [0, s_0]$,

$$-\frac{k_2 e_0 s}{K_M + s} \leq \dot{s} \leq -\frac{k_2 e_0 s}{K_M + s}\left(\frac{1}{1 + \max\limits_{s \in \ell} \varphi(s)}\right), \quad t_{\text{cross}} \leq t. \tag{3.38}$$

Since

$$\max_{0 \leq s \leq s_0} \varphi(s) = \frac{e_0}{K_M} = \varepsilon_{RS}, \tag{3.39}$$

we have:

$$-\frac{k_2 e_0 s}{K_M + s} \leq \dot{s} \leq -\frac{k_2 e_0 s}{K_M + s} \cdot \sum_{j=0}^{\infty} (-1)^j \varepsilon_{RS}^j, \quad t_{\text{cross}} \leq t, \quad \varepsilon_{RS} < 1. \tag{3.40}$$

This analysis has several significant consequences. First, we have established the sQSSA's validity depending on ε_{RS} without relying on non-dimensionalization, unlike traditional methods [see 71, 145, 163, among others]. This is advantageous due to the non-uniqueness of non-dimensionalization, which can lead to different "small parameters" [71, 145, 164]. Second, the bound (3.40) is more aesthetic than the estimate in Eilertsen et al. [51]:

$$-\frac{k_2 e_0 s}{K_M + s} \leq \dot{s} \leq -(1-m)\frac{k_2 e_0 s}{K_M + s}, \quad t_{\text{cross}} \leq t, \tag{3.41}$$

where "m" is the slope of the slow tangent vector, v^+.

Third, and this is somewhat unexpected, we gain more information about the approximation error of the flow on S_0^ε from $\varphi(s)$ than from m. From (3.37), the sQSSA approximates the flow on S_0^ε well in the regions where:

$$K_M e_0 \ll (K_M + s)^2, \tag{3.42}$$

even if $\varepsilon_{RS} \sim 1$. In particular, the sQSSA is a good approximation in regions where:

$$\max\{K_M, e_0\} \ll s. \tag{3.43}$$

This relates to the validity of the reverse quasi-steady-state approximation (rQSSA), which operates in high enzyme concentrations. To revisit the sQSSA's predominance, we examine the rQSSA's validity and investigate any overlap in their conditions.

3.4.2 Validity of the sQSSA for High Enzyme Concentrations

Equation (3.37) implies that the sQSSA is an excellent approximation to the flow on S_0^ε when $K_M \ll s$ and $e_0 \lesssim s$. In this region, the graphs of the c-nullcline and $\gamma(s)$ approach

their horizontal asymptote, $c = e_0$, reflected in $\varphi(s) \to 0$ as $s \to \infty$. The slow manifold, S_0^ε, is nearly horizontal implying $\dot{c} \approx 0$ on S_0^ε even for larger ε_{RS}.

Two singular perturbation scenarios lead to a nearly horizontal slow manifold: the sQSSA, where the slow manifold coalesces with the s-axis as $k_1 e_0 \to 0$, and the rQSSA, coinciding with small k_{-1} and small k_2:

$$\dot{s} = -k_1(e_0 - c)s + \varepsilon\widehat{k}_{-1}c, \tag{3.44a}$$

$$\dot{c} = \quad k_1(e_0 - c)s - \varepsilon\widehat{k}_{-1} - \varepsilon\widehat{k}_2. \tag{3.44b}$$

In the rQSSA scenario, the set of stationary points in the singular limit is not a submanifold of $\mathbb{R}^2$:

$$S_0 = S_0^{(1)} \cup S_0^{(2)}$$

where[8]

$$S_0^{(1)} = \{(s,c) \in \mathbb{R}^2_{\geq 0} : s = 0 \;\&\; 0 \leq c \leq e_0 - \zeta_1\}, \quad 0 < \zeta_1 < e_0, \tag{3.45a}$$

$$S_0^{(2)} = \{(s,c) \in \mathbb{R}^2_{\geq 0} : c = e_0 \;\&\; \zeta_2 \leq s \leq s_0\}, \qquad 0 < \zeta_2 < s_0. \tag{3.45b}$$

Classical Fenichel theory does not apply to the entire set, but it applies to specific compact submanifolds. The resulting Fenichel reduction via projection onto $TS_0^{(2)}$ is:

$$\dot{s} = -k_2 e_0, \tag{3.46a}$$

$$\dot{c} = 0. \tag{3.46b}$$

[8]The constants ζ_1, ζ_2 are introduced to enforce the compactness of $S_0^{(1)}$ and $S_0^{(2)}$.

Likewise, the Fenichel reduction obtained via projection onto $TS_0^{(1)}$ is:

$$\dot{s} = 0, \tag{3.47a}$$

$$\dot{c} = -k_2 c. \tag{3.47b}$$

In the rQSSA, trajectories initially approach the line $c = e_0$ and stay close until reaching the vicinity of the transcritical bifurcation point, $(s, c) = (0, e_0)$. Near this point, trajectories approach $(0, 0)$ as $t \to \infty$; however, the slow eigenvector in this scenario is nearly indistinguishable from the c-axis (and in fact aligns with the c-axis in the singular limit).

The rQSSA's long-time validity requires $\varepsilon_{RS}^{-1} \ll 1$, while the sQSSA's long-time validity requires $\varepsilon_{RS} \ll 1$ [42]. However, through comparison, the sQSSA with (3.46) reveals that it is practically indistinguishable from the rQSSA when $K_M \ll e_0 \lesssim s$. Thus, the sQSSA can approximate the flow on S_0^ε to the right of the bifurcation point, extending its validity beyond the Reich-Sel'kov criterion.

A natural question is: If we use the sQSSA to approximate the flow on S_0^ε to the right of the bifurcation instead of (3.46), how reliable is it in an asymptotic sense? Since S_0^ε lies within Γ_0 when $0 < \varepsilon$, we can get a rough answer. Assuming $\varepsilon_{RS}^{-1} \ll 1$ and considering $s \in [e_0, s_0]$, it follows from (3.37) that, for $t \geq t_{\text{cross}}$,

$$\begin{aligned}
-\frac{k_2 e_0 s}{K_M + s} \leq \dot{s} &\leq -\frac{k_2 e_0 s}{K_M + s}\left(\frac{1}{1 + \max\limits_{s\in[e_0,s_0]} \varphi(s)}\right) \\
&= -\frac{k_2 e_0 s}{K_M + s}\left(\frac{1}{1+\mu}\right), \quad \mu := \frac{\varepsilon_{RS}}{(1+\varepsilon_{RS})^2} < \varepsilon_{RS}^{-1} \\
&\leq -\frac{k_2 e_0 s}{K_M + s}\left(\frac{1}{1+\varepsilon_{RS}^{-1}}\right) \\
&= -\frac{k_2 e_0 s}{K_M + s} \cdot \sum_{j=0}^{\infty} (-1)^j \varepsilon_{RS}^{-j}, \quad 1 < \varepsilon_{RS}.
\end{aligned} \tag{3.48}$$

The bound (3.48) implies the sQSSA is a good approximation when $\varepsilon_{RS}^{-1} \ll 1$ and $e_0 \lesssim s$, improving as $e_0/s \to 0$.

While the sQSSA approximates the flow on S_0^ε well when $\varepsilon_{RS}^{-1} \ll 1$ and $e_0 \lesssim s$, it cannot be equipped with s_0 as the initial substrate concentration in the rQSSA scenario. If the initial condition is $(s, c)(0) = (s_0, 0)$, with $e_0 < s_0$, Fenichel theory dictates that the reduction (3.46) should be equipped with $(s, c) = (s_0 - e_0, e_0)$. This extends to the sQSSA if used to approximate the flow on S_0^ε to the right of the bifurcation point when $\varepsilon_{RS}^{-1} \ll 1$.

The existence of multiple accurate QSS reductions under the same conditions necessitates evaluating which QSSA is predominant. To truly verify the standard-QSS reduction's applicability, we must explore other QSS reductions near the sQSS curve in the phase-plane. As noted in Section 3.3, in scenarios like large k_{-1}, the sQSSA (3.5) and the slow product QSSA (3.7) are indistinguishable. To find the most accurate reduced system and its validity conditions, we next investigate the slow manifold's location relative to the (algebraic) variety that generates the slow product QSS reduction (3.7).

3.5 Trapping the Slow Manifold: The Slow Product QSSA

This section defines a new upper fence and anti-funnel for the slow manifold using the slow product QSS curve under specific parametric restrictions. We compare this to previous results to assess the approximation accuracy of the sQSSA (3.5) and the slow product QSSA (3.7), revisiting the Reich-Sel'kov condition and investigating whether it ensures the predominance of the sQSSA.

The slow product QSS reduction (3.7), derived from Fenichel theory for small k_2, corresponds to the QSS curve

$$c = \gamma_{SP}(s) := \frac{e_0 s}{K_S + s} - \frac{k_2 e_0 s}{k_1(K_S e_0 + (K_S + s)^2)}. \tag{3.49}$$

This closely resembles the curve $\gamma(s)$. In Section 3.4.1, we established that the slow manifold lies between $\gamma(s)$ and $\gamma_c(s)$. Now, we explore how the slow product QSS variety fits into

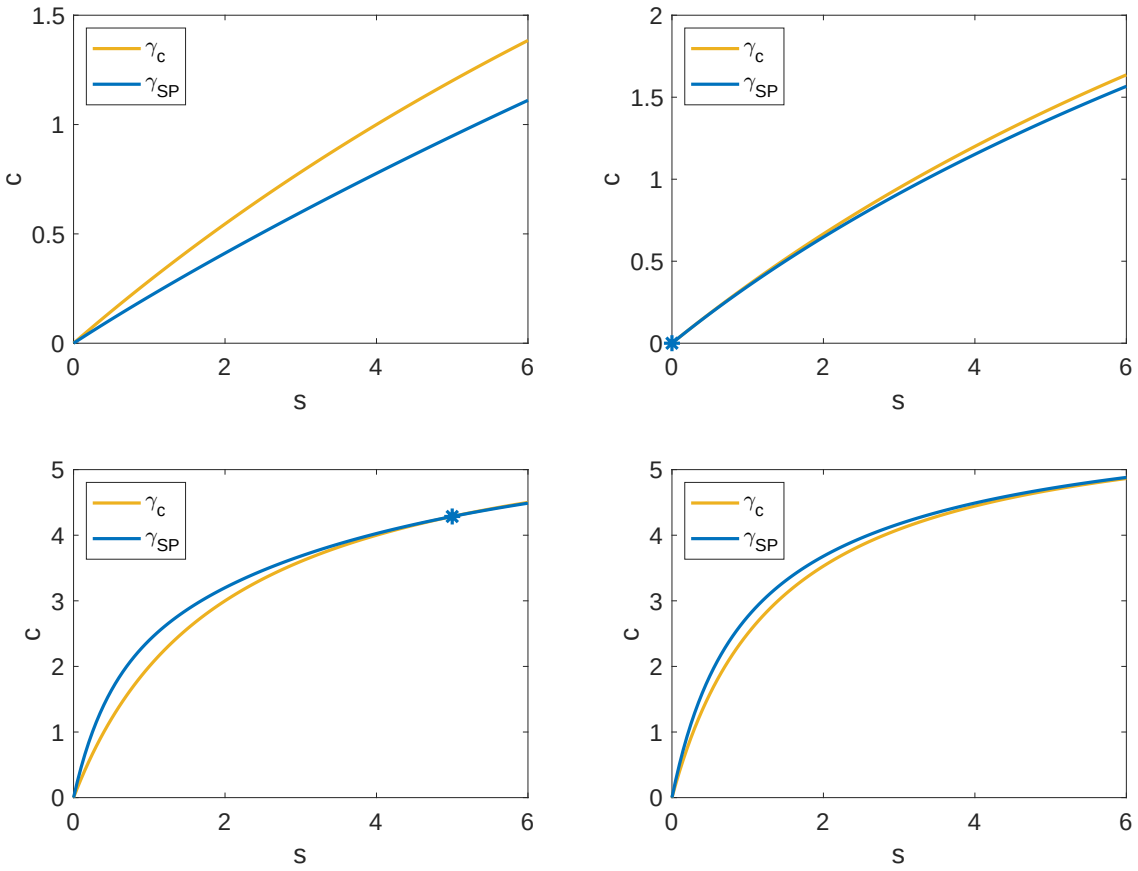


Figure 3.4: **The slow product QSS curve γ_{SP} lies above the horizontal nullcline γ_c for $s < s^*$ (3.50) in the phase plane.** The blue asterik denotes the intersection point at s^* (3.50). **Top Left ($k_2 > e_0 k_1$)**: $k_1 = 0.1$, $k_2 = 1$, $k_{-1} = 1$, $e_0 = 6$, $s^* = -4$. **Top Right ($k_2 = e_0 k_1$)**: $k_1 = 0.1$, $k_2 = 0.6$, $k_{-1} = 1$, $e_0 = 6$, $s^* = 0$. **Bottom Left ($k_2 < e_0 k_1$)**: $k_1 = 1$, $k_2 = 1$, $k_{-1} = 1$, $e_0 = 6$, $s^* = 5$. **Bottom Right ($k_2 < e_0 k_1$)**: $k_1 = 1$, $k_2 = 0.4$, $k_{-1} = 1$, $e_0 = 6$, $s^* = 14$.

these findings and the insights its location provides concerning the sQSSA's validity and predominance.

In the phase plane, $\gamma_{SP}(s)$ lies above the horizontal nullcline $\gamma_c(s)$ for $s < s^*$ where

$$s^* = \frac{k_{-1}e_0}{k_2} - \frac{k_{-1}}{k_1} = \frac{k_{-1}}{k_2}(e_0 - K) \tag{3.50}$$

is their intersection point for substrate concentration. We are primarily interested in cases where s^* is positive, placing the intersection within the first quadrant. When s^* is negative,

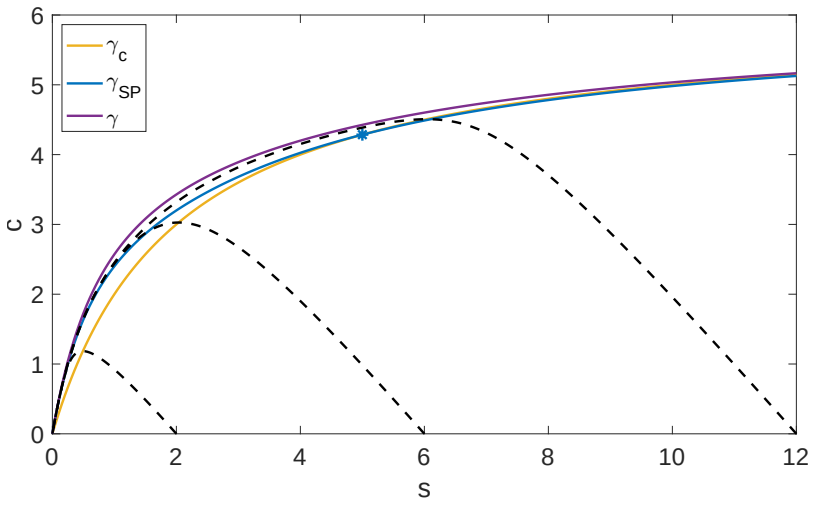


Figure 3.5: **The region between the slow product QSS curve $\gamma_{SP}(s)$ and the curve $\gamma(s)$ is a trapping region for solutions in the phase plane when** $e_0 < 8K_S$. The intersection point of γ_{SP} and γ_c, s^*, is denoted by the blue asterik. The numerical solutions to the mass action equations (3.1) for several initial conditions are denoted by black dashed lines. $s_0 \in \{2, 6, 12\}$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 1$, $k_2 = 1$.

or equivalently, when

$$e_0 < K, \tag{3.51}$$

the slow manifold crosses the sQSS variety, and the sQSSA is the best known reduction.

However, as e_0 increases, s^* shifts to the right. FIGURE 3.4 shows how the curves' positions change with parameter values. When $e_0 > K$, the intersection point lies in the first quadrant and $\gamma_{SP}(s) \geq \gamma_c(s)$ for $0 \leq s \leq s^*$. Also, $\gamma_{SP}(s) \leq \gamma(s)$ for all $s \geq 0$. Thus, the slow product QSS curve lies within Γ_0 for a significant portion of the phase plane. This raises the questions: Do solutions cross $\gamma_{SP}(s)$ after crossing $\gamma_c(s)$ under any parametric conditions? If so, can we locate the slow manifold more precisely?

Numerical evidence suggests this is possible, with solutions lying closer to the slow product QSS curve in the steady-state regime of the phase-plane (see Figure 3.5). Interestingly, the parametric conditions for this also ensure the slow-product QSS reduction's predominance over the sQSS reduction in the slow regime, aligning with our overall goal.

Our main results establish the conditions for positive invariance of the set bordered by

the graphs of $\gamma_{SP}(s)$ and $\gamma(s)$:

$$\Gamma_{SP} := \left\{(s,c) \in \mathbb{R}^2_{\geq 0} \ : \ \gamma_{SP}(s) \leq c \leq \gamma(s)\right\}, \tag{3.52}$$

and show that all solutions eventually cross $\gamma_{SP}(s)$ under those conditions.

Theorem 3.3. For a solution $(s,c)(t)$ to (3.2) with initial conditions $(s,c)(0) = (s_0, 0)$. If $e_0 < 8K_S$, then there exists a $t_{\text{cross_sp}} > 0$ at which the trajectory crosses $\gamma_{SP}(s)$, and

$$(s,c)(t) \in \Gamma_{SP}, \quad \forall\, t \geq t_{\text{cross_sp}}. \tag{3.53}$$

Proof. To establish positive invariance, we show that the vector field at the boundary curve γ_{SP} points towards Γ_{SP} when $e_0 < 8K_S$:

$$\dot{c}(s, \gamma_{SP}(s)) - \gamma'_{SP}(s)\dot{s}(s, \gamma_{SP}(s)) > 0 \tag{3.54}$$

We first compute the differential terms at γ_{SP}:

$$\begin{aligned}
\dot{c}(s, \gamma_{SP}(s)) &= k_1 e_0 s - k_1(K_M + s)\gamma_{SP}(s), && \text{(3.55a)}\\
&= k_1 e_0 s - k_1 e_0 s \frac{(K_M + s)}{(K_S + s)} + \frac{k_2 e_0 s (K_M + s)}{K_S e_0 + (K_S + s)^2}, && \text{(3.55b)}\\
&= -\frac{k_2 e_0 s}{(K_S + s)} + \frac{k_2 e_0 (K_M + s) s}{K_S e_0 + (K_S + s)^2}, && \text{(3.55c)}
\end{aligned}$$

and

$$\begin{aligned}
\dot{s}(s, \gamma_{SP}(s)) &= -k_1 e_0 s + k_1(K_S + s)\gamma_{SP}(s), && \text{(3.56a)}\\
&= -\frac{k_2 e_0 (K_S + s) s}{K_S e_0 + (K_S + s)^2}. && \text{(3.56b)}
\end{aligned}$$

The expression at $c = \gamma_{SP}(s)$ is:

$$\dot{c} - \gamma'_{SP}(s)\dot{s} = -\frac{k_2 e_0 s}{(K_S + s)} + \frac{k_2 e_0 (K_M + s)s}{K_S e_0 + (K_S + s)^2} + \frac{k_2 e_0 (K_S + s)s}{K_S e_0 + (K_S + s)^2}\gamma'_{SP}(s), \tag{3.57a}$$

$$= \frac{k_2 e_0 s}{(K_S + s)(K_S e_0 + (K_S + s)^2)}\big(-(K_S e_0 + (K_S + s)^2) + \tag{3.57b}$$

$$(K_M + s)(K_S + s) + (K_S + s)^2 \gamma'_{SP}(s)\big) \tag{3.57c}$$

Proving inequality (3.54) simplifies to showing:

$$-(K_S e_0 + (K_S + s)^2) + (K_M + s)(K_S + s) + (K_S + s)^2 \gamma'_{SP}(s) > 0. \tag{3.58}$$

This also proves that $\gamma_{SP}(s)$ is an upper fence. Substituting $\gamma'_{SP}(s)$:

$$\gamma'_{SP}(s) = \frac{K_S e_0}{(K_S + s)^2} - \frac{k_2 e_0}{k_1 (K_S e_0 + (K_S + s)^2)} + \frac{2k_2 e_0 (K_S + s)s}{k_1 (K_S e_0 + (K_S + s)^2)^2}, \tag{3.59a}$$

and simplifying, we obtain:

$$\frac{k_2}{k_1}(K_S + s) - \frac{k_2 e_0 (K_S + s)^2}{k_1 (K_S e_0 + (K_S + s)^2)} + \frac{2k_2 e_0 (K_S + s)^3 s}{k_1 (K_S e_0 + (K_S + s)^2)^2} \tag{3.60}$$

which is positive for $e_0 < 8K_S$. To show this, we further simplify the inequality to:

$$(K_S + s)^4 + e_0 (K_S + s)^3 - K_S e_0^2 s > 0, \tag{3.61a}$$

$$1 + \frac{e_0}{(K_S + s)} - \frac{K_S e_0^2 s}{(K_S + s)^4} > 0. \tag{3.61b}$$

Substituting $e_0 = \eta K_S$ and introducing $x = K_S/(K_S + s)$, we get:

$$1 + \frac{e_0}{(K_S + s)} - \frac{K_S e_0^2 s}{(K_S + s)^4} = 1 + \eta x - \eta^2 x^3 (1 - x) = \eta^2 x^4 - \eta^2 x^3 + \eta x + 1. \tag{3.62}$$

We want to show that $\eta^2 x^4 - \eta^2 x^3 + \eta x + 1 > 0$ for $\eta < 8$. Dividing by η^2 and factoring the

quartic function, we get:

$$x^4 - x^3 + \frac{1}{\eta}x + \frac{1}{\eta^2} = \left(x^2 + \bar{\eta}_+ x - \frac{1}{\eta}\right)\left(x^2 + \bar{\eta}_- x - \frac{1}{\eta}\right) \tag{3.63}$$

where

$$\bar{\eta}_\pm = -\frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{\eta - 8}{\eta}}.$$

This quartic function has no real roots when $\eta < 8$ and remains positive. It has 2 real repeated roots for $\eta = 8$ and 4 real roots when $\eta > 8$. Combining this with (3.29a) ensures solutions entering Γ_{SP} remain within. This completes the proof for the first part of Theorem 3.3.

We also show that solutions beginning outside Γ_{SP} eventually enter it. In contradiction, assume that a solution starting at $(s_0, 0)$ converges to the stationary point $(0, 0)$ without entering Γ_{SP}. Then, the slope of the tangent to $c = \gamma_{SP}(s)$ at $s = 0$ should be greater than m (3.18). However, we show that

$$\gamma'_{SP}(0) < m \tag{3.64}$$

always holds, where:

$$\gamma'_{SP}(0) = \frac{k_1 e_0}{k_{-1}} - \frac{k_2 e_0}{k_{-1}(K_S + e_0)}. \tag{3.65}$$

When $e_0 < K$, $\gamma'_{SP}(0) < \gamma'_c(0)$ holds, and Theorem 3.2 implies $\gamma'_c(0) < m$, demonstrating $\gamma'_{SP}(0) < m$. To prove the inequality for $e_0 > K$, recall:

$$m = \frac{1}{2k_{-1}}\left(-k_{-1} - k_2 + k_1 e_0 + \sqrt{(k_{-1} + k_2 + k_1 e_0)^2 - 4k_1 k_2 e_0}\right) \tag{3.66}$$

and

$$\gamma'_{SP}(0) = \frac{k_1 e_0}{k_{-1}} - \frac{k_2 k_1 e_0}{k_{-1}(k_1 e_0 + k_{-1})}. \tag{3.67}$$

Therefore, $m - \gamma'_{SP}(0)$ is:

$$m - \gamma'_{SP}(0) = \frac{-(k_{-1} + k_2 + k_1 e_0)}{2k_{-1}} + \frac{k_2 k_1 e_0}{k_{-1}(k_1 e_0 + k_{-1})} \tag{3.68a}$$

$$+ \frac{1}{2k_{-1}} \sqrt{(k_{-1} + k_2 + k_1 e_0)^2 - 4k_1 k_2 e_0}, \tag{3.68b}$$

$$= \frac{-(k_{-1} + k_1 e_0)}{2k_{-1}} + \frac{k_2(k_1 e_0 - k_{-1})}{2k_{-1}(k_1 e_0 + k_{-1})} + \frac{1}{2k_{-1}} \sqrt{(k_{-1} + k_2 + k_1 e_0)^2 - 4k_1 k_2 e_0}, \tag{3.68c}$$

$$= \frac{-(k_{-1} + k_1 e_0)}{2k_{-1}} + \frac{k_2(k_1 e_0 - k_{-1})}{2k_{-1}(k_1 e_0 + k_{-1})} \tag{3.68d}$$

$$+ \frac{1}{2k_{-1}} \sqrt{(k_{-1} + k_1 e_0)^2 + k_2^2 + 2k_2 k_{-1} - 2k_1 k_2 e_0}. \tag{3.68e}$$

Simplifying this expression, we want to show:

$$\frac{-(k_{-1} + k_1 e_0)}{2k_{-1}} + \frac{k_2(k_1 e_0 - k_{-1})}{2k_{-1}(k_1 e_0 + k_{-1})} + \frac{1}{2k_{-1}} \sqrt{(k_{-1} + k_1 e_0)^2 + k_2^2 - 2k_2(k_1 e_0 - k_{-1})} > 0 \tag{3.69}$$

or

$$\frac{1}{2k_{-1}} \sqrt{(k_{-1} + k_1 e_0)^2 + k_2^2 - 2k_2(k_1 e_0 - k_{-1})} > \frac{(k_{-1} + k_1 e_0)}{2k_{-1}} - \frac{k_2(k_1 e_0 - k_{-1})}{2k_{-1}(k_1 e_0 + k_{-1})}. \tag{3.70}$$

The above expression is equivalent to:

$$(k_1 e_0 + k_{-1}) \sqrt{(k_{-1} + k_1 e_0)^2 + k_2^2 - 2k_2(k_1 e_0 - k_{-1})} > (k_1 e_0 + k_{-1})^2 - k_2(k_1 e_0 - k_{-1}). \tag{3.71}$$

When $k_2 < k_1 e_0$, both sides are positive. Thus, we can prove that the above inequality is true by showing that the corresponding squared terms satisfy the same inequality. In simpler words, if $a, b > 0$ and $a^2 > b^2$, then $a > b$. Squaring both sides and simplifying, we can verify

that:

$$(k_1e_0 + k_{-1})^2((k_{-1} + k_1e_0)^2 + k_2^2 - 2k_2(k_1e_0 - k_{-1})) > \\ (k_1e_0 + k_{-1})^4 + k_2^2(k_1e_0 - k_{-1})^2 - 2k_2(k_1e_0 - k_{-1})(k_1e_0 + k_{-1})^2. \quad (3.72)$$

Hence, the inequality $m - \gamma'_{SP}(0) > 0$ holds when $k_2 < k_1e_0$.

Thus, our assumption is false, and a $t_{\text{cross_sp}} > 0$ exists when the trajectory crosses $\gamma_{SP}(s)$. The proof is complete by combining the two arguments. □

The primary benefit of establishing the positive invariance of Γ_{SP} is that it allows for a more precise localization of the slow manifold.

Theorem 3.4. If $e_0 < 8K_S$, then for the differential equation (3.13), Γ_{SP} is a positively invariant, narrowing anti-funnel within which there exists a unique slow manifold, S_0^ε.

Proof. Inequality (3.54) ensures $c = \gamma_{SP}(s)$ is a strong upper fence for (3.13) when $e_0 < 8K_S$. It is verifiable that $\gamma(s) > \gamma_{SP}(s)$ for $s > 0$ and:

$$\lim_{s\to\infty} |\gamma(s) - \gamma_{SP}(s)| = 0 \quad (3.73)$$

implying Γ_{SP} (3.28) is a narrowing anti-funnel for $e_0 < 8K_S$. Further,

$$\frac{\partial f}{\partial c}(c, s) \geq 0 \quad (3.74)$$

in Γ_{SP}, and the claim follows from Theorem 3.1. □

The slow manifold is contained within Γ_{SP} when $e_0 < 8K_S$. While this condition might seem restrictive, it connects directly to the sQSSA's validity analysis. We show that the slow manifold lies closer to the slow product QSS curve than the sQSS curve when $e_0 < 8K_S$. The remaining task is to analyze the implications of this location.

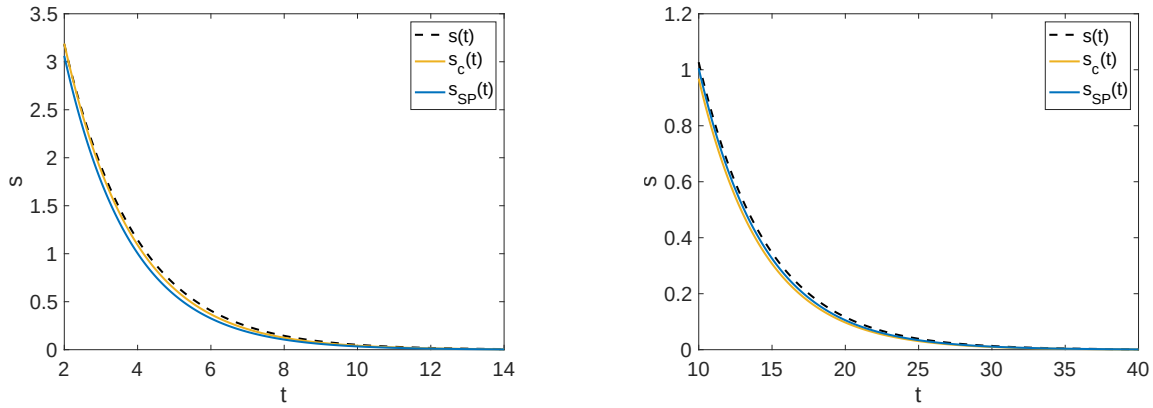


Figure 3.6: **The condition $e_0 < K$ ensures that the sQSSA provides a better approximation than the slow product QSSA under the Reich-Sel'kov condition.** $e_0 \ll K_M$ is satisfied in both panels. **Left** ($e_0 < K$): $s_0 = 9$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 100$, $k_2 = 10$. **Right** ($e_0 > K$): $s_0 = 9$, $c_0 = 0$, $e_0 = 6$, $k_1 = 1$, $k_{-1} = 100$, $k_2 = 4$. Substrate depletion over time is shown in both panels. The numerical solution to the mass action equations (3.1) is denoted by black dashed lines. $s_c(t)$ is the numerical solution to (3.5) and $s_{SP}(t)$ is the numerical solution to (3.7).

3.5.1 Revisiting the Reich-Sel'kov Condition

The Reich-Sel'kov qualifier ensures the sQSS reduction's accuracy, as validated in Section 3.4 using the novel anti-funnel (3.2). Our analysis shows that the sQSSA is the best known approximation to the slow manifold under most parametric scenarios. For instance, recall the intersection point (3.50). When $e_0 < K$, the intersection is inconsequential. Consequently, when the Reich-Sel'kov condition $e_0 \ll K_M$ is combined with $e_0 < K$, the sQSSA is the predominant approximation over the slow product QSSA.

What happens when $e_0 \ll K_M$, $e_0 \ll K_S$ (ensuring $e_0 < 8K_S$), but $e_0 > K$? Theorem 3.3 implies that trajectories lie closer to γ_{SP} than to γ_c for a significant portion of the slow dynamics (see FIGURE 3.5), and therefore

$$-\frac{k_2 e_0 s}{K_M + s} \leq -\frac{k_2 e_0 s(K_S + s)}{K_S e_0 + (K_S + s)^2} \leq \dot{s} \leq -\frac{k_2 e_0 s(K_M + s)}{K_M e_0 + (K_M + s)^2} \tag{3.75}$$

after the trajectories cross γ_{SP} and $s \leq s^*$. Moreover, we know that the slope m of the linear approximation is greater than the slope of the tangent to γ_{SP} at the origin (3.64) and

$\gamma'_{SP}(0) > \gamma'_c(0)$ whenever $e_0 > K$. Thus, the approximation errors near the stationary point follow:

$$m - \gamma'_{SP}(0) < m - \gamma'_c(0) \tag{3.76}$$

when $e_0 > K$. Consequently, complementing the Reich-Sel'kov condition with the stronger condition $e_0 < K$ is necessary to guarantee that the sQSS is indeed the predominant quasi-steady-state approximation of the MM reaction mechanism, and this may be relevant in the context of the inverse problem where parameters are estimated at low substrate concentration [168]: If $e_0 > K$, using the slow product QSS variety as an approximation for small s is beneficial. FIGURE 3.6 illustrates the numerical substrate depletion curves. The restrictive condition $e_0 < K$ is necessary to ensure the sQSS most accurately approximates the MM dynamics for small substrate concentrations.

3.6 Discussion

The analysis in this chapter addresses a critical gap in the understanding of the quasi-steady-state approximation for the MM reaction mechanism. While the sQSSA is a widely used tool in enzyme kinetics, the precise conditions for its validity have remained a topic of ongoing investigation. We conducted a thorough phase-plane analysis of the MM mass-action kinetics using the theory of fences and anti-funnels, a powerful technique for analyzing the long-time behavior of dynamical systems.

Our analysis led to the identification of new positively invariant sets that contain the slow manifold, S_0^ε, in the substrate-complex phase plane. These sets provide valuable insights into the dynamics of the system and allow for a more precise characterization of the sQSSA's accuracy. As a result, we obtained improved bounds on the estimation error of various QSS approximations, including the standard, reverse, and slow product formation approximations, in the slow regime.

Significantly, we have demonstrated that the commonly accepted qualifier for the validity of the sQSSA, the Reich-Sel'kov condition ($e_0 \ll K_M$), does not guarantee that the sQSSA

is the predominant reduction. Predominance necessitates a more restrictive condition

$$e_0 \ll K, \tag{3.77}$$

where K is the Van Slyke-Cullen constant. This finding challenges the traditional understanding of the sQSSA's validity and highlights the importance of considering both accuracy and predominance when evaluating QSS approximations.

Our findings have broader implications for the application of QSS approximations in various fields. By providing a more refined understanding of the sQSSA's validity, our work can guide researchers in selecting the most appropriate and accurate reduction for their specific needs. This is particularly crucial in areas such as quantitative biology and pharmacology, where accurate model reduction is essential for understanding complex biological processes and designing effective therapeutic interventions.

Future research could extend our analysis by considering more complex reaction mechanisms or incorporating additional factors that might influence the sQSSA's validity. Further investigation of the error associated with the trajectory's approach to the slow manifold is also warranted. By addressing these open questions, we can continue to refine our understanding of QSS approximations and enhance their utility in diverse scientific disciplines.

CHAPTER 4

Gaining Hyperbolicity: A Blow-up Approach to Understanding Quasi-Steady-State Approximations

4.1 Introduction

For a perturbed form of a dynamical system,

$$\dot{x} = h(x, \varepsilon), \quad x \in \mathbb{R}^n,\ h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n,\ 0 < \varepsilon \ll 1, \tag{4.1}$$

where "$\dot{x}$" denotes differentiation of x with respect to time and ε denotes a small perturbation parameter, the 'singular' limit corresponds to the layer problem at $\varepsilon = 0$,

$$\dot{x} = h(x, 0). \tag{4.2}$$

For a two-scale system (corresponding to fast and slow variables), the layer problem essentially describes the fast dynamics of the system. Fenichel theory, on the other hand, approximates the slow flow away from the singular limit i.e., for $\varepsilon > 0$ and proposes the reduced system that estimates the slow dynamics of the original (full) system [56, 55]. Underscoring its practical utility to real-world processes, The Tikhonov-Fenichel Parameter Value (TFPV) theory establishes the required conditions that justify the use of quasi-steady-state reductions [63, 65, 66].

Put together, the Fenichel and TFPV theories necessitate biochemical parametric condi-

tions that ensure that the set of stationary points associated with the layer problem,

$$\Gamma := \{x \in \mathbb{R}^n : h(x, 0) = 0\}, \tag{4.3}$$

is *normally hyperbolic*. Generally speaking, if Γ is a non-empty submanifold of $\mathbb{R}^n$ with dimension $k < n$ i.e.,

$$\text{rank } \mathrm{D}_1 h(x, 0) = n - k \qquad \forall\, x \in \Gamma, \tag{4.4}$$

where $\mathrm{D}_1 h(x, 0)$ refers to the Jacobian of $h(x, \varepsilon)$ with respect to x at $\varepsilon = 0$, then (4.1) is a singularly perturbed system. Here, the set of stationary points (4.3) is called the critical manifold. Further, the critical manifold is normally hyperbolic if the $n - k$ non-zero eigenvalues of $\mathrm{D}_1 h(x, 0)$ are strictly bounded away from the imaginary axis for all $x \in \Gamma$. In such cases, Γ *survives* small perturbations and becomes a slow invariant manifold, Γ^ε, for $\varepsilon > 0$. Fenichel theory dictates that the non-trivial flow on Γ^ε can be approximated by projecting the perturbation onto the tangent space of Γ, wherein the explicit definition of the projection is made plausible by the presence of normal hyperbolicity.[1]

But what exactly happens when normal hyperbolicity is lost?

Several omnipresent physical phenomenon such as bifurcation points [48], center manifolds, oscillations [69, 120], folds and canard points [105, 169] are all accompanied with the loss of normal hyperbolicity. Not only do the Fenichel theory and TFPV theory fail to apply at such non-hyperbolic singularities, but this also introduces additional multi-layered challenges. First, how do we extend GSPT to find and validate (modified) reduced systems when normal hyperbolicity is lost? Second, if we instead utilize other reduced systems which do not result from GSPT, how do we mathematically justify their application and then analyze their validity as well? These questions remain open for rigorous investigation beyond the scope of classical ODE analysis tools [46].

While Fenichel theory does extend to non-hyperbolic or bifurcating critical manifolds,

[1]We urge readers to refer to exhaustive texts on GSPT and TFPV theory for details on derivation, convergence and examples of quasi-steady-state reductions [106, 184, 43]. See Chapter 1 for a brief overview.

the theory mostly pertains to the normal forms of the aforementioned bifurcations [103, 104]. However, bifurcations occurring in real world systems often contain imperfections that can distort the structure of the bifurcation. These imperfections essentially act like a "perturbation on top of a perturbation" in which the normal form of the bifurcation is perturbed. Such a perturbation can result in a loss of symmetry, or even break up intersecting critical manifolds [67, 46]. Systems with imperfections usually require a case-by-case analysis.

Thus, in this chapter, we address these challenges for an important enzymatic mechanism called the Intermolecular Autocatalytic Zymogen Activation (IAZA) mechanism:

$$\mathrm{Z} + \mathrm{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_2} 2\,\mathrm{E} + \mathrm{W} \tag{4.5}$$

where Z is a zymogen, E is an active enzyme, C is an intermediary complex and W is a peptide.

We employ the blow-up method, a powerful geometrical approach that transforms the dynamics within critical regions of the phase-plane [75, 86]. This technique allows for a detailed examination of the behavior in these regions by regaining normal hyperbolicity and providing insights that traditional methods often miss. See Section 4.2 for additional motivation. Notably, to the best of our knowledge, the blow-up method has not been extensively applied to non-prototypical systems.

Using a combination of standard analysis and the blow-up method, we perform a thorough analysis of the validity of the most-commonly employed QSS reductions for the IAZA mechanism. We derive qualifying conditions that provide the quantitative accuracy of QSS reductions near a non-hyperbolic stationary point and comment on the surprising effectiveness of a heuristic reduction that does not attribute to GSPT. The aim is to not just analyze the IAZA mechanism under these difficult scenarios but also establish the required mathematical toolbox that can be applied to a broad range of mechanisms.

The chapter is structured as follows. In Section 4.2, we set up the motivation for using

the blow-up method by detailing the limitations of the Fenichel theory pertinent to its real-world applications. In Section 4.3, we provide a synopsis of the blow-up method and the main terminology required by our work. This review is kept high-level for accessibility. In Section 4.4, we review the IAZA mechanism and its phase-plane dynamics in detail. Following that, we focus on the perturbation scenario of small k_1 and small k_2 for the IAZA system revealing a non-hyperbolic singularity in Section 4.5. We perform a blow-up analysis of the system to investigate the validity of the reduced systems near the bifurcation point. We analyze the deviations from the normal form theory in Section 4.6. The current work is concluded in Section 4.7.

4.2 Limitations of Fenichel theory

Despite extensive studies on the theoretical and practical utility of quasi-steady-state reductions, several gaps remain where Fenichel theory is limited in real-world applications. In this section, we outline some of these challenges, map them to the IAZA mechanism, and relate them to the overarching objective of addressing these limitations using the blow-up method.

4.2.1 Loss of Normal Hyperbolicity

As mentioned in Section 4.1, Fenichel's theory is applicable only for the Tikhonov-Fenichel parameter values (TFPVs) which produce normally hyperbolic critical manifolds for a perturbed system. Thus, the crucial question arises - What happens in cases where normal hyperbolicity is lost? Specifically, for compact critical sets, Γ, several mechanisms involve isolated singular points $x^* \in \Gamma$ where

$$\text{rank } \mathrm{D}_1 h(x^*, 0) < n - k \tag{4.6}$$

implying that the oblique projection matrix onto the tangent space of Γ cannot be defined. See Chapter 1 for more details.

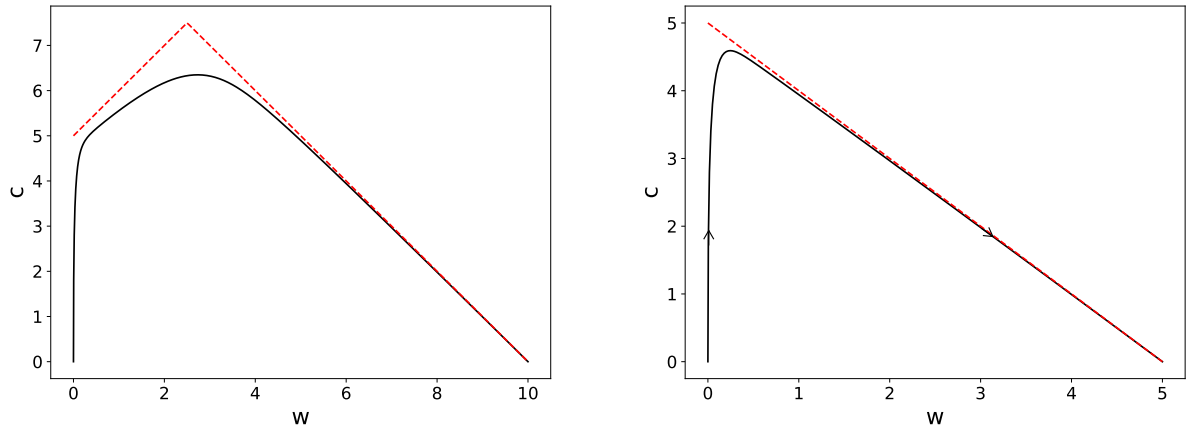


Figure 4.1: **The movement of solution trajectories in the phase-plane when $z_0 > e_0$ and $z_0 < e_0$.** The solid black curve is the numerical solution to the full IAZA ODE model (4.7) and the dashed red curve is the numerical solution to the reverse-QSSA (4.44) in the (w, c) phase-plane. **Left**: The parameter values are $k_1 = 10, k_{-1} = 1, k_2 = 1, z_0 = 10, e_0 = 5$. **Right**: The parameter values are $k_1 = 10, k_{-1} = 1, k_2 = 1, z_0 = 5, e_0 = 10$.

For the system of ordinary differential equations that models (4.5),

$$\dot{w} = k_2 c, \tag{4.7a}$$

$$\dot{c} = k_1(e_0 + w - c)(z_0 - w - c) - (k_{-1} + k_2)c, \tag{4.7b}$$

w and c denote the concentrations of Peptide and Complex respectively and z_0 and e_0 denote the initial concentrations of Zymogen and Enzyme respectively.[2] Now, consider the perturbation scenario with the parameter values $\pi^* := (k_1, \varepsilon k_{-1}, \varepsilon k_2, z_0, e_0)^T$, $0 < \varepsilon \ll 1$, pertaining to small k_{-1} and small k_2.[3] The layer problem,

$$\begin{pmatrix} \dot{w} \\ \dot{c} \end{pmatrix} = \begin{pmatrix} 0 \\ k_1(e_0 + w - c)(z_0 - w - c) \end{pmatrix}, \tag{4.8}$$

[2]These equations can be obtained using law of mass action and appropriate conservation laws.

[3]The parameter space for the IAZA system is a 5-tuple, $\pi = (k_1, k_{-1}, k_2, z_0, e_0)^T$ where T denotes the vector transpose.

corresponds to the critical set

$$\Gamma_0 = \Gamma_1 \cup \Gamma_2 = \{(w, c) \in \mathbb{R}^2_{\geq 0} : c = e_0 + w\} \cup \{(w, c) \in \mathbb{R}^2_{\geq 0} : c = z_0 - w\}. \tag{4.9}$$

When $z_0 > e_0$, we observe that S_0 loses normal hyperbolicity at the (biologically relevant) intersection point of Γ_1 and Γ_2,

$$(w^*, c^*) = ((z_0 - e_0)/2, E_T/2), \tag{4.10}$$

where E_T is the sum of initial concentrations of zymogen and enzyme, $E_T = z_0 + e_0$. The point (w^*, c^*) acts as a transcritical bifurcation point. Consequently, the projection of the perturbation onto the tangent space of Γ_0 is infeasible and Fenichel theory breaks down. One approach that has been employed in the past involves avoiding a small neighborhood of the intersection point, resulting in a "modified" Fenichel reduction, commonly referred to as the reverse quasi-steady-state reduction [163, 162]. What remains lacking is a thorough understanding of the quasi-steady-state reduction in the vicinity of the bifurcation point.

This warrants the extension of classical reduction theory to systems that admit such singularities [103, 104]. Possible alternative methods can include some straightforward and popular techniques such as linearization or center manifold theory. However, as we will observe shortly, loss of normal hyperbolicity is often accompanied with *nilpotency* of the critical set, i.e., all eigenvalues of the Jacobian, $\mathrm{D}_1 h(x, 0)$, are trivial at these points. This renders linearization ineffective. The center manifold is the entire state space, making the center manifold theory inconsequential as well. Thus, we turn to a more complex yet useful *blow-up method* instead.

If a singular point in the critical set is *nilpotent*, i.e.,

$$\text{rank } \mathrm{D}h_1(x^*, 0) = 0 \;\text{ for some }\; x^* \in \Gamma, \tag{4.11}$$

the blow-up method is used to geometrically desingularize the vector field in its neighborhood [75, 37]. For our work, we focus on the IAZA mechanism to understand the application of blow-up method in cases where classical reduction theory fails.

4.2.2 Quantitative Understanding of ε

As mentioned in Chapter 1, another limitation of Fenichel theory is the lack of quantitative understanding of the small perturbation parameter ε. While several efforts have been made for the systematic derivation of plausible qualifiers, we encounter roadblocks for systems with nilpotent singular points.

Methods such as timescale separation or eigenvalue ratios are inconsequential (all eigenvalues are equal to 0) due to the singular nature of such systems [46]. Through the application of blow-up method, we also extract quantitative insight into the validity conditions of reduced systems when normal hyperbolicity is lost.

As we will review in Section 4.5, heuristic strategies which cannot be attributed to GSPT or Fenichel theory have been widely studied in the literature to obtain reductions for systems under such singularity constraints [61, 115, 187]. Thus, it becomes highly crucial to (a) identify a rigorous mathematical justification for popular heuristic reductions and (b) investigate the parametric qualifiers necessary for the validity of Fenichel reductions over heuristic reductions. This will be another focus of our results.

4.3 The Toolbox: Blow-up, Quasi-homogeneity and Coordinate Charts

In this section, we present a brief literature review of the blow-up method and associated theory such as quasi-homogeneity of vector fields and coordinate charts. These will serve as the main tools for the phase-plane analysis of the IAZA system that follows.

4.3.1 The Blow-up Map

The blow-up method was developed by Hironaka [76, 75] for applications in the field of algebraic geometry and was later extended to the study of multiple time-scale systems

by Dumortier and Roussarie [38]. Other excellent works focus on the extension of the blow-up method for commonly-found singularities such as canards, fold points, bifurcation points, relaxation oscillations and center manifolds [38, 103, 104, 105, 119]. We shadow the exposition by Kuehn [106] to introduce the basic terminology of the method.

Consider the dynamical system

$$\dot{x} = H(x), \quad x \in \mathbb{R}^n,\ H : \mathbb{R}^n \to \mathbb{R}^n \tag{4.12}$$

where H is smooth and $H(0) = 0$.[4] Assume that all the eigenvalues of the Jacobian, $\mathrm{D}H(x)$ vanish at $x = 0$, leading to loss of normal hyperbolicity at the 'nilpotent' origin. Thus, the qualitative behavior of the solutions near the equilibrium point is unclear. To circumnavigate this, we use the blow-up method here. Although the blow-up method allows various forms of coordinate transformations, let us begin by defining a general polar blow-up as it is the most intuitive to understand.

Definition 4.1. For system (4.12), the polar coordinate transformation is given by

$$\varphi(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, r) = (r\bar{x}_1, r\bar{x}_2, \ldots, r\bar{x}_n), \quad \varphi : S^{n-1} \times I \to \mathbb{R}^n \tag{4.13}$$

where S^{n-1} denotes the $(n-1)^{th}$-sphere such that $\sum_{i=1}^{n} \bar{x}_i^2 = 1$ and $I \subseteq \mathbb{R}$ such that $0 \in I$. The polar blow-up $\bar{H}$ of the vector field H is[5]

$$\bar{H}(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, r) := ((\mathrm{D}\varphi)^{-1} \circ H \circ \varphi)(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, r). \tag{4.14}$$

Here, the Jacobian of φ, $\mathrm{D}\varphi$, is computed with respect to the hyperspherical coordinates. The coordinate transformation essentially replaces the origin in the original phase-space with a sphere in the blown-up space for $r = 0$ in order to gain suitable hyperbolicity. Since φ

[4]Here, x is a vector that includes the perturbation parameter ε.

[5]The blow-up map is typically defined by the function φ. However, it is actually the function φ^{-1} that maps the original space to the blown-up space. This is just a notational nuance followed by the existing literature.

is a diffeomorphism everywhere except $\{r = 0\}$, the dynamics away from the equilibrium point remain intact in the blown-up space. Further, Dumortier [37] shows the continuous extension of $\bar{H}$ to $\{r = 0\}$.

Algebraically, the blow-up would imply the following schematics:

$$\begin{array}{ccc} T(S^{n-1} \times I) & \xrightarrow{\mathrm{D}\varphi} & T(\mathbb{R}^n) \\ {\scriptstyle \bar{H}}\uparrow & & \uparrow{\scriptstyle H} \\ S^{n-1} \times I & \xrightarrow{\varphi} & \mathbb{R}^n \end{array}$$

where $T(\cdot)$ denotes the tangent space.

It might as well happen that $\bar{H}$ is trivial for $r = 0$. In order to retrieve non-trivial dynamics of the blown-up system, we remove the singularity at $r = 0$ by dividing $\bar{H}$ by the maximum possible degree of r: $\bar{\mathcal{H}} = \bar{H}/r^k$. Now, $\bar{\mathcal{H}}$ is non-zero at $r = 0$ and closely approximates the flow denoted by $\bar{H}$ for small $r > 0$ (and equivalently the flow in H away from the singularity) as well. The vector field $\bar{\mathcal{H}}$ is called the 'rescaled' version of the blown-up vector field. Solutions to the ODE system with $\bar{H}$ are the same as solutions to the system with $\bar{\mathcal{H}}$ up to a time rescaling or time reparameterization. Geometrically, time rescaling does not change the shape of trajectories in the phase plane, but it changes the speed or parameterization at which trajectories are traversed. It is this rescaled vector field that will desingularize the original vector field. Please refer to [106] for detailed text.

As one might guess, it is not ensured that a single blow-up transformation results in complete desingularization of the vector field. Further, some vector fields might not admit an interpretation in terms of angular variables (unlike, for example, flow in a circular pipe) and might require repeated blow-ups to gain hyperbolicity. There is, however, one class of vector fields which adopts a natural setting for the blow-up method to work well and avoid such cumbersome computations: quasi-homogenenous vector fields.

4.3.2 Quasi-Homogeneity of Vector Fields

Definition 4.2. A function $h : \mathbb{R}^n \to \mathbb{R}$ is called *quasi-homogeneous* if

$$h(r^{a_1}x_1, r^{a_2}x_2, \ldots, r^{a_n}x_n) = r^k\, h(x_1, x_2, \ldots, x_n) \tag{4.15}$$

where $r \in \mathbb{R}$, $(a_1, a_2, \ldots, a_n) \in \mathbb{N}^n$ is called the *type* of quasi-homogeneity and k is called the degree of quasi-homogeneity.

Definition 4.3. A vector field H is called *quasi-homogeneous* of degree $k+1$ and type $(a_1, a_2, \ldots, a_n)$ if its ith function component, h_i, is quasi-homogeneous of degree $k + a_i$ and type $(a_1, a_2, \ldots, a_n)$.

For example, consider a non-trivial ODE system given by

$$\dot{x_1} = h_1(x_1, x_2) = x_1^2 + x_1x_2,$$
$$\dot{x_2} = h_2(x_1, x_2) = x_1x_2 + x_2^2.$$

We can verify that $h_1(x_1, x_2)$ and $h_2(x_1, x_2)$ are quasi-homogeneous functions of type (1,1) and degree 2. Thus, the vector field $H = (h_1(x_1, x_2), h_2(x_1, x_2))$ is also quasi-homogeneous of type (1,1) and degree 2.

An obvious way to write a blow-up map for a quasi-homogeneous vector field follows.

Definition 4.4. For system (4.12), define the general polar transformation

$$\varphi(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, r) = (r^{a_1}\bar{x}_1, r^{a_2}\bar{x}_2, \ldots, r^{a_n}\bar{x}_n), \quad \varphi : S^{n-1} \times I \to \mathbb{R}^n \tag{4.17}$$

where S^{n-1} denotes the $(n-1)^{th}$-sphere such that $\sum_{i=1}^{n} \bar{x}_i^2 = 1$ and $I \subseteq \mathbb{R}$ such that $0 \in I$. The quasi-homogeneous blow-up $\bar{H}$ of the vector field H is given by (4.14).

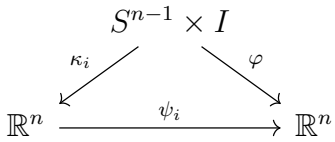
Again, the blow-up vector field can be rescaled by dividing the maximum possible common factor of r : $\bar{\mathcal{H}} = \bar{H}/r^k$. For our analysis, we will be taking advantage of the quasi-

homogeneity of the vector field associated with the IAZA mechanism to define the appropriate blow-up maps.

In the application of the blow-up method, a tangible problem that we face is how to define $(\mathrm{D}\varphi)^{-1}$. From (4.17), it might appear that φ maps $\mathbb{R}^{n+1}$ to $\mathbb{R}^n$. However, note that the blown-up space $B := S^{n-1} \times I$ is a smooth, n-dimensional submanifold of R^{n+1}.[6] Thus, there are open sets surrounding every point in B that are diffeomorphic to $\mathbb{R}^n$. We can assign an n-dimensional coordinate system to these open sets of B. Therefore, what we want to do is select suitable open subsets, U_i, of $S^{n-1} \times \mathbb{R}$ and a corresponding diffeomorphism, κ_i such that

$$\kappa_i : U_i \subset B \subset \mathbb{R}^{n+1} \to \mathbb{R}^n. \tag{4.18}$$

The ordered pair, (U_i, κ_i), is called a coordinate chart. The first entry, U_i, is an open set of the manifold B; the second entry, κ_i, is a diffeomorphism that assigns a unique coordinate in $\mathbb{R}^n$ to every point belonging to $U_i \subset B$. In practice for chemical mechanisms, we define the coordinate charts in a way so they cover the relevant parts of the phase-space for analysis, for instance the parts where solution trajectories enter or exit. Most manifolds cannot be covered with a single coordinate chart. Generally, an atlas, $A := \{U_i, \kappa_i\}$, which is a collection of open sets, U_i, and diffeomorphims, κ_i, is required to completely cover the entire manifold. Once we define an appropriate coordinate chart, (U_i, κ_i), we want to define a local coordinate transformation, $\psi_i : \mathbb{R}^n \to \mathbb{R}^n$, that satisfies the commutative diagram:



The coordinate charts are not defined in a standardized or a unique way. It truly depends on the system of equations and the analysis that we want to perform. An effective way to cover the blown-up space would be to utilize Cartesian coordinates in the directions of the blown-up variables. This is also known as a directional blowup and is widely used to define simpler coordinate charts for phase-plane analysis.

[6]The claim follows from $\sum_{i=1}^{n} \bar{x}_i^2 = 1$ and the rank theorem.

4.3.3 Defining Coordinate Charts using Directional Blow-up(s)

Recall the blown-up state variables $(\bar{x}_1, \ldots, \bar{x}_n, r)$. The fundamental idea for a coordinate chart in the direction of a coordinate, say $\bar{x}_i$, is that it follows[7]

$$\kappa_i(\bar{x}_1, \ldots, \bar{x}_n, r) = \left(\frac{\bar{x}_1}{\bar{x}_i^{b_1}}, \ldots, \frac{\bar{x}_{i-1}}{\bar{x}_i^{b_{i-1}}}, \frac{\bar{x}_{i+1}}{\bar{x}_i^{b_{i+1}}}, \ldots, \frac{\bar{x}_n}{\bar{x}_i^{b_n}}, \frac{r}{\bar{x}_i^{b_i}}, \right) := (\varkappa_1, \ldots, \varkappa_{i-1}, \varkappa_{i+1}, \ldots, \varkappa_n, \mathfrak{r}). \tag{4.19}$$

The chart κ_i is well-defined away from the set $\{(\bar{x}_1, \ldots, \bar{x}_n) \in S^{n-1} : \bar{x}_i = 0\}$. Here, $b_j \in \mathbb{R}$ are constant coefficients that are strategically chosen to ensure the commutativity $\varphi = \psi_i \circ \kappa_i$. Essentially, using the appropriate coefficients, we should be able to eliminate the coordinate $\bar{x}_i$ using the function composition.

Note that we are now working with weighted affine coordinates instead of "polar" coordinates. The motivation for this switch is straightforward. Using commutativity, we obtain the original coordinates using the map ψ_i on local Cartesian coordinates as

$$\psi_i(\varkappa_1, \ldots, \varkappa_{i-1}, \varkappa_{i+1}, \ldots, \varkappa_n, \mathfrak{r}) = (\mathfrak{r}^{a_1}\varkappa_1, \ldots, \mathfrak{r}^{a_{i-1}}\varkappa_{i-1}, \mathfrak{r}^{a_i}, \mathfrak{r}^{a_{i+1}}\varkappa_{i+1}, \ldots, \mathfrak{r}^{a_n}\varkappa_n) \tag{4.20a}$$

$$:= (x_1, \ldots, x_n), \tag{4.20b}$$

where the domain for ψ_i is the same as the range of κ_i. The map, ψ_i, denotes the weighted directional blowup and is a simpler transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. Even though multiple coordinate charts are required to cover the blown-up space, breaking down the computations into local coordinates using the intermediary directional blowups is usually preferred.

Finally, the last step is to obtain local vector fields defined on these coordinate charts individually. After performing a local analysis, these are combined together through transition maps to obtain a holistic picture of the overall dynamics. For two coordinate charts, κ_i and κ_j, the transition map from κ_i to κ_j is the function $\kappa_j \circ \kappa_i^{-1}$. Readers are encouraged to find more details in [106]. In our work, we are primarily concerned with the analysis of

[7]Note that the subscripting of the input and output of κ is kept consistent to improve readability.

local dynamical behavior around singular points and will not be relying on transition maps.

4.4 The IAZA Reaction Mechanism

We begin with a quick review of the phase-plane analysis of the IAZA model. This allows us to locate the solution trajectories in the phase-plane relative to the nullclines and provide mathematical context for the commonly-used reduced models.

4.4.1 Dynamics in (w, c) Phase-Plane

The system in (w, c) coordinates (4.7) gives rise to three nullclines in the phase-plane, $c = 0$ and $c = h^{\pm}(w)$ such that

$$h^{-}(w) = \frac{1}{2}((E_T + K_M) - \sqrt{(E_T + K_M)^2 - 4(z_0 - w)(e_0 + w)}), \tag{4.21a}$$

$$h^{+}(w) = \frac{1}{2}((E_T + K_M) + \sqrt{(E_T + K_M)^2 - 4(z_0 - w)(e_0 + w)}). \tag{4.21b}$$

The equilibrium point $(w, c)^1 = (-e_0, 0)$ is a saddle point and the equilibrium point $(w, c)^2 = (z_0, 0)$ is a stable node, see Figure 4.2.

For experimentally relevant initial conditions, the solution trajectories start at the origin and remain in the first quadrant in the (w, c) phase plane.[8] The c-nullcline, $c = h^{-}(w)$, provides an upper bound for the complex concentration for the system in positive time. Thus, the apex or the peak of the nullcline curve becomes a value of high interest to successfully implement any reduction strategies for the solution trajectories.

Consider the two cases depicted in Figure 4.2. When $e_0 \leq z_0$, basic calculus tells us that the peak of the curve, $c = h^{-}(w)$, occurs at $w = (z_0 - e_0)/2$ and the complex concentration is bounded as

$$c \leq h^{-}((z_0 - e_0)/2); \qquad t \geq 0. \tag{4.22}$$

On the other hand, when $e_0 > z_0$, the maximum value of the nullcline curve occurs at $w = 0$

[8] Generally, initial conditions for the IAZA mechanism satisfy: $z(0) = z_0, e(0) = e_0, w(0) = 0$ and $c(0) = 0$.

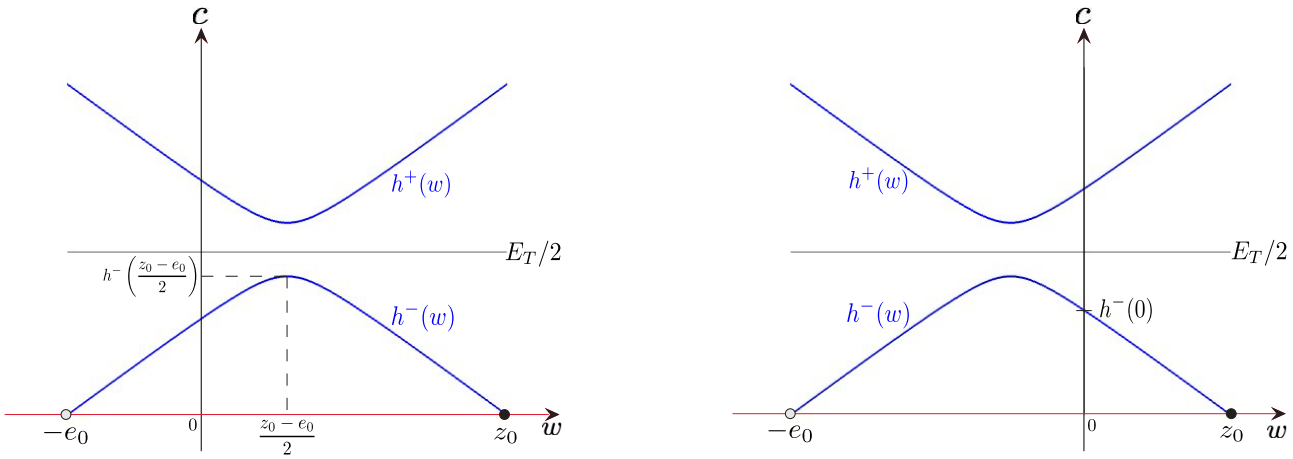


Figure 4.2: The phase plane geometry of the IAZA mechanism in (w, c) coordinates contains three nullclines. The w-nullcline is given by the w-axis depicted in red. The c-nullcline is given by (4.21) depicted by the two blue curves. The curve that intersects with the w-axis is $c = h^-(w)$. The two equilibrium points are at $(-e_0, 0)$ and $(z_0, 0)$ which are saddle and stable points respectively. **Left**: $e_0 \leq z_0$. The complex concentration of a solution trajectory is bounded above by $h^-((z_0 - e_0)/2)$. **Right**: $e_0 > z_0$. The complex concentration of a solution trajectory is bounded above by $h^-(0)$.

which implies

$$c \leq h^-(0); \qquad t \geq 0. \tag{4.23}$$

As we go on to explore the various quasi-steady-state reductions for the IAZA mechanism, it will be evident that the study of nullclines holds high significance in finding accurate approximations and their validity.

4.4.2 Solution Trajectories in the Phase-Plane

As discussed in Chapter 3, we gain valuable insight into reduced models by examining the location of solution trajectories in the phase plane. Several QSS approximations that do not derive from Fenichel theory can still be justified - albeit on a case-by-case basis - using tools from phase-plane analysis. Before proceeding to the singularly perturbed case with a non-hyperbolic point, we further investigate biologically relevant solution trajectories in the IAZA phase plane.

In particular, we present the following result.

Theorem 4.1. For the system (4.7), solution trajectories that initialize on the w-axis cross the nullcline curve, $h^-(w)$, exactly once for some $w > (z_0 - e_0)/2$.

Proof. We divide the proof in two steps.

1. We show that the solution trajectories that initialize on the w-axis cross the nullcline curve at least once for $w > (z_0 - e_0)/2$ before they approach the stable stationary point at $(z_0, 0)$.

2. We show that the trajectories cross the nullcline exactly once.

The stationary point $(z_0, 0)$ is a stable node with corresponding eigenvalues

$$\lambda_\pm = \frac{k_1}{2}\left(-(K_M + E_T) + \sqrt{(K_M + E_T)^2 - 4KE_T}\right) \tag{4.24}$$

where $\lambda_- < \lambda_+ < 0$. The trajectories approach the stationary point along the slow eigenvector with slope m:

$$m = \frac{\lambda_+}{k_2} = \frac{1}{2K}\left(-(K_M + E_T) + \sqrt{(K_M + E_T)^2 - 4KE_T}\right). \tag{4.25}$$

Secondly, the slope of the tangent to the curve $c = h^-(w)$ at $(z_0, 0)$ is

$$\alpha = \frac{\mathrm{d}h^-(w)}{\mathrm{d}w}(z_0, 0) = -\frac{E_T}{K_M + E_T}. \tag{4.26}$$

Note that both slopes m and α are negative. Thus, if $\alpha - m > 0$, then the solution trajectories approach the stationary point above the nullcline curve. This inequality can be easily shown. Consider the expression with $K = k_2/k_1$:

$$\alpha - m = -\frac{E_T}{K_M + E_T} + \frac{K_M + E_T}{2K} - \frac{1}{2K}\left(\sqrt{(K_M + E_T)^2 - 4KE_T}\right), \tag{4.27a}$$

$$= \frac{(K_M + E_T)^2 - 2KE_T}{2K(K_M + E_T)} - \frac{1}{2K}\left(\sqrt{(K_M + E_T)^2 - 4KE_T}\right). \tag{4.27b}$$

We want to show

$$\frac{(K_M + E_T)^2 - 2KE_T}{2K(K_M + E_T)} - \frac{1}{2K}\left(\sqrt{(K_M + E_T)^2 - 4KE_T}\right) > 0 \tag{4.28}$$

or

$$(K_M + E_T)^2 - 2KE_T > (K_M + E_T)\sqrt{(K_M + E_T)^2 - 4KE_T}. \tag{4.29}$$

Squaring both sides, we want to prove the inequality

$$(K_M + E_T)^4 + 4(KE_T)^2 - 4KE_T(K_M + E_T)^2 > (K_M + E_T)^4 - 4KE_T(K_M + E_T)^2 \tag{4.30a}$$

which always holds true. If $a, b > 0$ and $a^2 > b^2$, then $a > b$. Thus, the solutions lie above the nullcline curve as they approach $(z_0, 0)$ suggesting that they intersect the nullcline at least once before approaching the stationary point.

To find the values of w that permit this intersection in the phase plane, we utilize the direction of the vector field at the nullcline curve.

Given system (4.7), we compute the dynamics at the curve $c = h^-(w)$:

$$\dot{w}(w, h^-(w)) = k_2 h^-(w), \tag{4.31}$$

$$\dot{c}(w, h^-(w)) = 0. \tag{4.32}$$

The vector field for (4.7) points towards the w-axis for $w < (z_0 - e_0)/2$ and points away from the w-axis for $w > (z_0 - e_0)/2$. This is because

$$\dot{c}(w, h^-(w)) - \frac{\mathrm{d}h^-(w)}{\mathrm{d}w}\dot{w}(w, h^-(w)) = \left(\frac{(-z_0 + e_0 + 2w)}{\sqrt{(K_M + E_T)^2 - 4(z_0 - w)(e_0 + w)}}\right) k_2 h^-(w) \tag{4.33}$$

implying

$$\dot{c}(w, h^-(w)) - \frac{\mathrm{d}h^-(w)}{\mathrm{d}w}\dot{w}(w, h^-(w)) < 0 \quad \text{for} \quad w < \frac{z_0 - e_0}{2}, \tag{4.34}$$

$$\dot{c}(w, h^-(w)) - \frac{\mathrm{d}h^-(w)}{\mathrm{d}w}\dot{w}(w, h^-(w)) > 0 \quad \text{for} \quad w > \frac{z_0 - e_0}{2}. \tag{4.35}$$

Thus, the solution trajectories that initialize on the w-axis i.e., below the nullcline curve, can only intersect it for $w > (z_0 - e_0)/2$.

Finally, we aim to show that the trajectories cross the nullcline exactly once for some $w > (z_0 - e_0)/2$. The set

$$\Delta = \left\{ w > \frac{z_0 - e_0}{2} \quad \& \quad h^-(w) \leq c \leq z_0 - w \right\} \tag{4.36}$$

is positively invariant such that solutions that enter this set do not leave the set as $t \to \infty$. For this, we only need to show

$$\dot{c}(w, z_0 - w) - (-1)\dot{w}(w, z_0 - w) < 0 \quad \text{for} \quad \frac{z_0 - e_0}{2} < w < z_0 \tag{4.37}$$

since the derivative of the curve $c = z_0 - w$ with respect to w is -1. This is straighforward by computing the dynamics at the curve $c = z_0 - w$,

$$\dot{c}(w, z_0 - w) + \dot{w}(w, z_0 - w) = -(k_{-1} + k_2)(z_0 - w) + k_2(z_0 - w) \tag{4.38}$$

which is always negative for $w < z_0$. Thus, combining this with inequality (4.35), the set Δ is positively invariant. We already established that the solution trajectories cross the nullcline at least once in the range $(z_0 - e_0)/2 < w < z_0$. The positive invariance concludes that the solutions crosses the nullcline exactly once and remains above it in the phase-plane as it approaches the stationary point $(z_0, 0)$ for $t \to \infty$. □

By definition of the nullcline, the crossing occurs at a point where $\mathrm{d}c/\mathrm{d}t = 0$. We will refer

to this crossing point as the peak point in the phase plane. This corresponds to the point of maximum complex concentration. Applications in enzyme kinetics rely on the limiting rate, $(k_2E_T)/2$, representing the hypothetical scenario in which all enzyme molecules bind with zymogen molecules to form peptide. This outcome is rarely achieved in real-world practice. Instead, the observed maximum complex concentration, denoted by c^{peak}, provides a more realistic and experimentally viable estimate of this limiting rate. The limiting rate is critical for quantifying enzyme efficiency under near-saturation conditions and for calibrating kinetic parameters in experimental setups [160]. It enables more accurate predictions of reaction throughput in physiological and synthetic systems.

It is worth highlighting is that the solutions can never cross the curve $c = z_0 - w$. This is ensured by positive invariance shown in Theorem 4.1 and, more fundamentally, by the conservation law: $\dot{w} + \dot{c} + \dot{z} = 0$. As we discuss next, both the curves $c = h^{-}(w)$ and $c = z_0 - w$ are instrumental in the reduction of the IAZA system.

4.5 Loss of Normal Hyperbolicity: Dynamic Transcritical Bifurcation Point

For small k_{-1} and small k_2, the IAZA mechanism admits a perturbation scenario wherein normal hyperbolicity is lost at a transcritical bifurcation point. As a result, we can't obtain an accurate quasi-steady-state reduction or its justificaion near the singularity using Fenichel theory. Further, other reductions which do not attribute to GSPT have also been utilized in real-world applications. Thus, we employ the blow-up method to systematically gain adequate hyperbolicity and analyze the dynamics in the blown-up desingularized space for a complete study of the validity of these approximations.

We present the perturbation scenario in detail and recall the modified Fenichel reduction for this case - the reverse quasi-steady-state reduction. We also present a brief literature review on the classical reduction popularly employed in place of the reverse quasi-steady-state reductions. The classical reduction does not derive from Fenichel theory and surprisingly

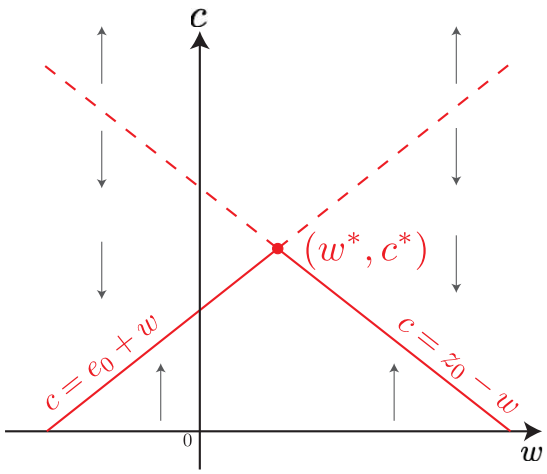


Figure 4.3: **The** (w, c) **phase-plane inhabitates a dynamic transcritical bifurcation point at the intersection of** Γ_1 **and** Γ_2 (4.9) **given by** (w^*, c^*) (4.41). The solid red lines denote the stable submanifolds and the dashed red lines denote unstable submanifolds.

provides suitable approximation close to the bifurcation point as well. We address this rigorously in our analysis.

4.5.1 Reduced Systems in the Presence of Bifurcation

As introduced in Section 4.2, the perturbation scenario of small k_{-1} and k_2 is,

$$\begin{pmatrix} \dot{w} \\ \dot{c} \end{pmatrix} = \begin{pmatrix} 0 \\ k_1(e_0 + w - c)(z_0 - w - c) \end{pmatrix} + \varepsilon \begin{pmatrix} k_2 c \\ -(k_{-1} + k_2)c \end{pmatrix}, \tag{4.39}$$

where parameters are scaled as $k_{-1} \to \varepsilon k_{-1}$ and $k_2 \to \varepsilon k_2$ with $\varepsilon \ll 1$.[9] The critical set $\Gamma_0 = \Gamma_1 \cup \Gamma_2$:

$$\Gamma_1 = \{(w, c) \in \mathbb{R}^2_{\geq 0} : c = e_0 + w\}, \tag{4.40a}$$

$$\Gamma_2 = \{(w, c) \in \mathbb{R}^2_{\geq 0} : c = z_0 - w\}. \tag{4.40b}$$

[9]The variable names are overwritten for simplicity.

loses normal hyperbolicity at the intersection point,[10]

$$(w^*, c^*) = ((z_0 - e_0)/2, E_T/2). \tag{4.41}$$

The intersection of Γ_1 and Γ_2 is a dynamic transcritical bifurcation point which depends on the initial conditions of the system. The two branches of the critical set exchange stability at the bifurcation point and the solution trajectories initially approach Γ_1 before finally approaching Γ_2, see Figure 4.3.

Due to the loss of normal hyperbolicity at (w^*, c^*), Fenichel theory cannot be applied to the entirety of the critical set. See Chapter 1 for details on the requirements and derivation of Fenichel reductions.

A work-around is found by ignoring a small neighborhood around the bifurcation point and observing that the two attracting branches

$$\Gamma_1^* := \{(w, c) \in \mathbb{R}^2_{\geq 0} : c = e_0 + w,\ 0 \leq w \leq w^* - \mu,\ 0 < \mu < w^*\}, \tag{4.42}$$

$$\Gamma_2^* := \{(w, c) \in \mathbb{R}^2_{\geq 0} : c = z_0 - w,\ w^* + \kappa \leq w \leq z_0,\ 0 < \kappa\}, \tag{4.43}$$

are compact and normally hyperbolic submanifolds of dimension 1. We apply Fenichel theory to the two disjoint branches separately and obtain a piecewise QSS reduction

$$\dot{w} = k_2(e_0 + w), \qquad 0 \leq w < w^*, \tag{4.44a}$$

$$\dot{w} = k_2(z_0 - w), \qquad w^* < w \leq z_0, \tag{4.44b}$$

known as the autocatalytic equivalent of the reverse quasi-steady-state approximation (rQSSA) [163, 162, 43]. It is essential to note that the existence of the bifurcation point is only biophysically relevant in the case $z_0 > e_0$. For initial values such that $e_0 > z_0$, the QSS reduction simplifies to $\dot{w} = k_2(z_0 - w)$ resulting from projection onto Γ_2^* only. Figure

[10] Here, rank $\mathrm{D}h(w, c) = 1$ for $(w, c) \in \Gamma_0/(w^*, c^*)$ and rank $\mathrm{D}h(w, c) = 0$ for $(w, c) = (w^*, c^*)$.

4.1 shows the numerical solution of the IAZA system (4.7) and the reverse QSSA (4.44) for the different cases of initial conditions, z_0 and e_0.[11]

Various works emphasize the success of the reverse-QSS curve in approximating the system dynamics for high enzyme concentrations in both the deterministic [162, 43] and stochastic regimes [48]. However, a missing piece of the puzzle is that the reverse-QSSA is modeled by ignoring the phase space near the bifurcation point.

To overcome the limitations of Fenichel theory, experimental applications of chemical kinetic theory generally utilize reduced systems which do not directly result from GSPT. In our case, a heuristic approach to obtain a *classical* quasi-steady-state reduction assumes that the complex concentration changes very slowly i.e., remains approximately constant, after a rapid accumulation over an initial transient time. Thus, for the majority of the reaction, the complex maintains a steady state. The classical QSS reduction for Peptide accumulation is

$$\dot{w} = k_2 h^-(w) \tag{4.45}$$

where $h^-(w)$ is the c-nullcline (4.21a). Since the above reduction strategy does not directly align with Fenichel theory, the reduction (4.45) does not correspond to a restricted parameter regime close to a specific TFPV in the parameter space. Unlike the QSS reductions obtained from Fenichel theory, research suggests that the classical QSS approximation can be reliably employed for applications such as parameter estimation for a wide range of biochemical parametric conditions. Wu et al. [187] provide evidence for accurate numerical results by the classical reduction specifically for the IAZA mechanism. Several other works highlight the success of such (heuristic) nullcline reductions in applications spanning but not limited to enzyme kinetics [134, 133, 177] and gene regulatory networks [10, 100] among other fields of mathematical biology.

The universality of a classical reduction seems to be overly simplistic yet advantageous. Limited work has been done on obtaining sound mathematical understanding of when and

[11] Moving forward, the figures only implement initial conditions such that $z_0 > e_0$.

why these reductions work as valid approximations for chemical mechanisms. Of course, these heuristic reductions work in our favor if they also provide accurate approximations near points where Fenichel theory fails i.e., near singularities. The missing aspect is a thorough justification for using classical reduction instead of Fenichel reduction in such cases. This motivates our work on the IAZA mechanism with singularities and investigating the success of the classical nullcline reduction (4.45).

4.5.2 Blow-up Analysis near the Bifurcation Point

To study the dynamics of systems in presence of singularities, one option is to eliminate (not ignore) the singularities and then use standard analysis tools. The blow-up method is specifically designed for the purpose of handling nilpotent non-hyperbolic singularities such as the transcritical bifurcation point (4.41) for IAZA mechanism.

To proceed, we first shift the bifurcation point to the origin. Substitute $c = \widehat{c} + E_T/2$ and $w = \widehat{w} + (z_0 - e_0)/2$ and note that the mass-action equations simplify to

$$\begin{pmatrix} \dot{\widehat{w}} \\ \dot{\widehat{c}} \\ \dot{\varepsilon} \end{pmatrix} = h(\widehat{w}, \widehat{c}, \varepsilon) + \varepsilon \widehat{c}, \varepsilon) = \begin{pmatrix} 0 \\ k_1(\widehat{c}^2 - \widehat{w}^2) \\ 0 \end{pmatrix} + \varepsilon \begin{pmatrix} k_2(\widehat{c} + E_T/2) \\ -(k_{-1} + k_2)(\widehat{c} + E_T/2) \\ 0 \end{pmatrix} \tag{4.46}$$

with shifted initial concentrations $\widehat{w}(0) = -(z_0 - e_0)/2$ and $\widehat{c}(0) = -E_T/2$. The equilibrium at the origin $(\widehat{w}, \widehat{c}, \varepsilon) = (0, 0, 0)$ corresponds to the bifurcation point where the Jacobian of the unperturbed system, $\mathrm{D}h(\widehat{w}, \widehat{c}, \varepsilon)$, only has eigenvalues equal to zero at the origin. Thus, the bifurcation point is a nilpotent equilibrium point and the blow-up method allows us to define a map from the original coordinates $(\widehat{w}, \widehat{c}, \varepsilon)$ to a different coordinate set for which the singularity no longer exists.

Let $B := S^2 \times \mathbb{R}$ and $B_0 := S^2 \times [0, r_0]$ for some $r_0 > 0$. To define the appropriate blow-up map, we leverage the *quasi-homogeneity* of the vector field at hand (see B.1 for the

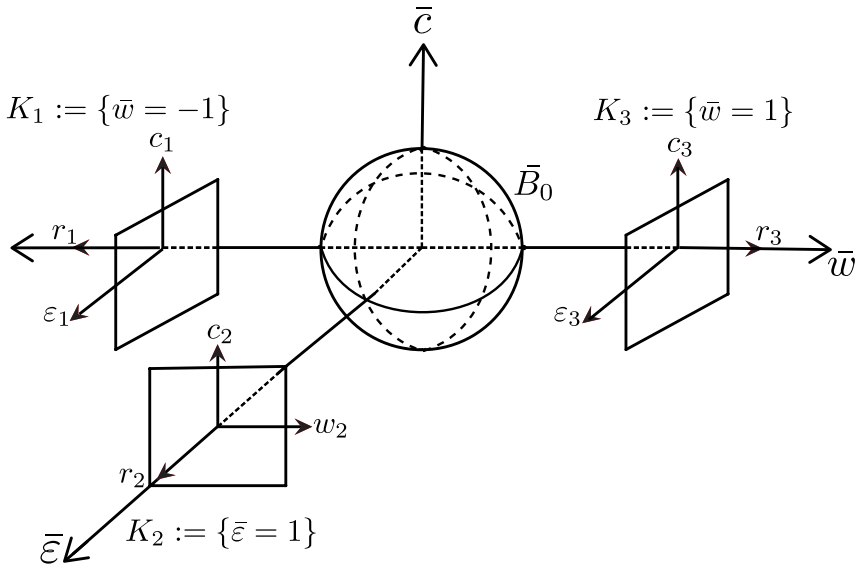


Figure 4.4: **The blown-up space can be divided into 3 coordinate charts.** The blown-up space is defined for coordinates $(\bar{w}, \bar{c}, \bar{\varepsilon}, r)$ and corresponds to the space $S^2 \times [0, r_0]$. Chart K_1 covers the region where trajectories enter. Chart K_2 covers a small neighborhood of the bifurcation point at the origin. Chart K_3 covers the region where trajectories exit. The origin in (w, c, ε) is mapped to $\bar{B}_0 := S^2 \times \{0\}$ for $r = 0$.

derivation). For system (4.46), we define the map as

$$\varphi(\bar{w}, \bar{c}, \bar{\varepsilon}, r) := (r\bar{w}, r\bar{c}, r^2\bar{\varepsilon}) = (\widehat{w}, \widehat{c}, \varepsilon), \quad \varphi : B_0 \to \mathbb{R}^3. \tag{4.47}$$

As one might guess, it is too cumbersome to analyze the system dynamics over the entire blow-up space B_0. Thus, we define intermediary *directional* blow-ups in the directions of the blown-up coordinates [103] by setting one of them equal to ± 1 at a time. For this, consider the shifted system (4.46) and note that the solution trajectories follow the curve $\{\widehat{c} = \widehat{w}\}$ until near the bifurcation point at the origin and then finally follows the curve $\{\widehat{c} = -\widehat{w}\}$. Thus, the parts of the blown-up space that we are interested in can be divided into three coordinate charts K_1, K_2 and K_3 that cover the directions $\bar{w}$, $\bar{\varepsilon}$ and $-\bar{w}$ respectively. Refer to Figure 4.4 for visualization. The direction $\bar{w}$ will cover the region that the trajectories enter, $\bar{\varepsilon}$ is for the region near the bifurcation point and finally the direction $-\bar{w}$ will demarcate

the region where the trajectories exit. Once the analysis is performed in the individual coordinate charts, we can *blow-down*[12] the analysis from each of the coordinate charts and study the overall dynamics of the original system by using transition maps between the charts.

Since we are only interested in the neighborhood of the bifurcation point, our analysis doesn't necessarily require and is rather hindered by the complexity of the blown-up vector fields on all 3 coordinate charts. For dynamics near the bifurcation point, we focus on the second coordinate chart and define local rectangular coordinates in the neighborhood of the origin,

$$K_2 := \quad \widehat{w} = r_2 w_2, \qquad \widehat{c} = r_2 c_2, \qquad \varepsilon = r_2^2, \tag{4.48}$$

through a directional blow-up in the direction of $\bar{\varepsilon}$.[13] This is popularly known as the rescaling chart since r_2 is fixed ($\dot{\varepsilon} = 0$) and the coordinate transformation is essentially a rescaling of $\widehat{w}$ and $\widehat{c}$. We urge readers to peruse B.2 for details on obtaining the coordinate charts through the intermediary directional blow ups and rectangular coordinates.

System (4.46) after applying the rescaling transformation (4.48) becomes

$$\dot{w_2} = k_2 r_2 \left(r_2 c_2 + E_T/2\right), \tag{4.49a}$$

$$\dot{c_2} = k_1 r_2 (c_2^2 - w_2^2) - r_2 (k_{-1} + k_2) \left(r_2 c_2 + E_T/2\right), \tag{4.49b}$$

$$\dot{r_2} = 0. \tag{4.49c}$$

Note that $r_2 = 0$ is a removable singularity here and we can desingularize the flow by

[12] *Blow-down* refers to the application of the coordinate chart transformations in the reverse order to obtain the dynamics in the original coordinates.

[13] The subscript 2 indicates the second coordinate chart. The first and third coordinate charts are detailed in B.2.

dividing the vector field by r_2 to obtain

$$\dot{w_2} = k_2(E_T/2) + \mathcal{O}(r_2), \tag{4.50a}$$

$$\dot{c_2} = k_1(c_2^2 - w_2^2) - (k_{-1} + k_2)(E_T/2) + \mathcal{O}(r_2), \tag{4.50b}$$

$$\dot{r_2} = 0, \tag{4.50c}$$

where $\mathcal{O}(r_2)$ is a first-order term in r_2.[14] This is now a classical problem in asymptotic analysis. We can find a power series $c_2(w_2, r_2)$ to approximate a trajectory that is an extension of the slow attracting manifold Γ_1^* into a small neighborhood of the bifurcation point:

$$c_2(w_2, r_2) = c_2(w_2, 0) + r_2 \frac{\partial c_2}{\partial r_2}(w_2, 0) + \mathrm{O}(r_2^2). \tag{4.51}$$

In fact, the zeroth-order term of the above trajectory is the solution to the one-dimensional differential equation

$$\frac{\mathrm{d}c_2}{\mathrm{d}w_2} = \frac{2k_1(c_2^2 - w_2^2) - (k_{-1} + k_2)E_T}{k_2 E_T}. \tag{4.52}$$

This is a Riccati differential equation whose solutions can be expressed in terms of special functions [1]. We present the following novel results to obtain the appropriate solution for our system.

Proposition 4.1. The general solution to the differential equation (4.52) is

$$f(w_2) = -w_2 - \sqrt{2\mu}\frac{(-\beta_1 D_{-b+1}(\phi(w_2)) + \beta_2 D_{-b+1}(-\phi(w_2)))}{\beta_1 D_{-b}(\phi(w_2)) + \beta_2 D_{-b}(-\phi(w_2))} \tag{4.53}$$

where $\mu = k_2 E_T/2k_1$, $b = (k_{-1} + 2k_2)/(2k_2)$, $\phi(w_2) = \sqrt{\frac{2}{\mu}} w_2$ and β_1 and β_2 are constants. Here, $D_\nu(z)$ is a parabolic cylinder function as defined by Whittaker and Watson [186].

Proof. We write the equation in the form of the generalized parabolic cylinder differential

[14] Here, the desingularization step is also equivalent to the time rescaling, $t' = r_2 t$, where t' denotes the new time. We rewrite the notation to avoid confusion.

equation $y'' + (az^2 + bz + c)y = 0$ [1, Chapter 19].

Let $\mu = k_2 E_T / 2k_1$ and $v = c_2/\mu$, then

$$\frac{\mathrm{d}v}{\mathrm{d}w_2} = v^2 - \frac{1}{\mu}\left(\frac{K_M}{K} + \frac{w_2^2}{\mu}\right) \tag{4.54}$$

Further, substitute $v = -\left(\frac{1}{u}\right)\frac{\mathrm{d}u}{\mathrm{d}w_2}$ to obtain

$$\frac{\mathrm{d}^2 u}{\mathrm{d}w_2^2} - \frac{1}{\mu}\left(\frac{K_M}{K} + \frac{w_2^2}{\mu}\right) u = 0. \tag{4.55}$$

The solutions of this equation can be expressed in terms of parabolic cylinder functions. To convert this into the standard form, substitute

$$w_2 = \sqrt{\frac{\mu}{2}} z$$

to obtain

$$\frac{\mathrm{d}^2 u}{\mathrm{d}z^2} - \left(\frac{K_M}{2K} + \frac{1}{4}z^2\right) u = 0. \tag{4.56}$$

We compare this to the standard form of a parabolic cylinder function [1, Equation 19.1.2]

$$y'' - \left(\frac{1}{4}x^2 + a\right) y = 0.$$

The solutions are given by

$$U(a, x) = D_{-a-\frac{1}{2}}(x), \tag{4.57a}$$

$$V(a, x) = \frac{1}{\pi}\Gamma\left(\frac{1}{2} + a\right)\left\{\sin(\pi a)\, D_{-a-\frac{1}{2}}(x) + D_{-a-\frac{1}{2}}(-x)\right\} \tag{4.57b}$$

where $D_\nu(z)$ is a parabolic cylinder function as defined by Whittaker and Whatson [186].

Thus, the general solution in our case is given parabolic cylinder functions for

$$a = \frac{K_M}{2K} \tag{4.58}$$

as

$$u(z) = \beta_1 D_{-b}(z) + \beta_2 D_{-b}(-z), \tag{4.59}$$

where

$$b = a + \frac{1}{2}, \tag{4.60a}$$

$$\beta_1 = \alpha_1 + \frac{\alpha_2}{\pi}\Gamma\left(\frac{1}{2} + a\right)\sin(\pi a), \tag{4.60b}$$

$$\beta_2 = \frac{\alpha_2}{\pi}\Gamma\left(\frac{1}{2} + a\right). \tag{4.60c}$$

We can retrieve the solution in terms of c_2 and w_2 by substituting the variables back. To find the general solution of (4.52), we first find $\mathrm{d}u/\mathrm{d}w_2$. In terms of z,

$$\frac{\mathrm{d}}{\mathrm{d}z}u(z) = \beta_1 D'_{-b}(z) - \beta_2 D'_{-b}(-z). \tag{4.61}$$

Parabolic cylinder functions, $D_\nu(z)$, follow the recurrence relations

$$D_{\nu+1}(z) - zD_\nu(z) + \nu D_{\nu-1}(z) = 0, \tag{4.62a}$$

$$D'_\nu(z) + \frac{z}{2}D_\nu(z) - \nu D_{\nu-1}(z) = 0. \tag{4.62b}$$

Therefore,

$$D'_{-b}(z) = -\frac{z}{2}D_{-b}(z) - bD_{-b-1}(z), \tag{4.63a}$$

$$D'_{-b}(-z) = \frac{z}{2}D_{-b}(-z) - bD_{-b-1}(-z). \tag{4.63b}$$

Substituting in (4.61),

$$\frac{\mathrm{d}u}{\mathrm{d}z} = \beta_1\left(-\frac{z}{2}D_{-b}(z) - bD_{-b-1}(z)\right) - \beta_2\left(\frac{z}{2}D_{-b}(-z) - bD_{-b-1}(-z)\right) \tag{4.64}$$

$$= -\frac{z}{2}\left(\beta_1 D_{-b}(z) + \beta_2 D_{-b}(-z)\right) + \left(-\beta_1 b D_{-b-1}(z) + \beta_2 b D_{-b-1}(-z)\right). \tag{4.65}$$

Using recurrence relation (4.62a) with $\nu = -b$,

$$-bD_{-b-1}(z) = \quad zD_{-b}(z) - D_{-b+1}(z), \tag{4.66a}$$

$$-bD_{-b-1}(-z) = -zD_{-b}(-z) - D_{-b+1}(-z), \tag{4.66b}$$

which yields

$$\begin{aligned}\frac{\mathrm{d}u}{\mathrm{d}z} &= -\frac{z}{2}\left(\beta_1 D_{-b}(z) + \beta_2 D_{-b}(-z)\right) + z\left(\beta_1 D_{-b}(z) + \beta_2 D_{-b}(-z)\right) \\ &\qquad + \left(-\beta_1 D_{-b+1}(z) + \beta_2 D_{-b+1}(-z)\right), \\ &= \frac{z}{2}u(z) + \left(-\beta_1 D_{-b+1}(z) + \beta_2 D_{-b+1}(-z)\right).\end{aligned}$$

In terms of w_2,

$$u(w_2) = \beta_1 D_{-b}\left(\sqrt{\frac{2}{\mu}}w_2\right) + \beta_2 D_{-b}\left(-\sqrt{\frac{2}{\mu}}w_2\right) \tag{4.67a}$$

$$\frac{\mathrm{d}u}{\mathrm{d}w_2} = \frac{w_2}{\mu}u(w_2) + \sqrt{\frac{2}{\mu}}\left(-\beta_1 D_{-b+1}\left(\sqrt{\frac{2}{\mu}}w_2\right) + \beta_2 D_{-b+1}\left(-\sqrt{\frac{2}{\mu}}w_2\right)\right) \tag{4.67b}$$

Recall $v = -(1/u)(\mathrm{d}u/\mathrm{d}w_2)$ and $c_2 = \mu v$ which yields

$$c_2 := f(w_2) = -w_2 - \sqrt{2\mu}\frac{\left(-\beta_1 D_{-b+1}\left(\sqrt{\frac{2}{\mu}}w_2\right) + \beta_2 D_{-b+1}\left(-\sqrt{\frac{2}{\mu}}w_2\right)\right)}{\beta_1 D_{-b}\left(\sqrt{\frac{2}{\mu}}w_2\right) + \beta_2 D_{-b}\left(-\sqrt{\frac{2}{\mu}}w_2\right)}. \tag{4.68}$$

□

The general solution (4.53) exhibits various phase-plane behaviors that are beyond the

focus of our analysis for the IAZA mechanism. We are primarily interested in the behavior of the solution as $w_2 \to \infty$, $w_2 \to -\infty$, and $w_2 \to 0$. Even though the original coordinate $\widehat{w}$ is physically relevant only for finite values, the scaling $\widehat{w} = w_2 r_2$ implies that the blown-up coordinate w_2 approaches $\pm\infty$ as r_2 approaches 0.

In particular, we observe that solution trajectories in the (w_2, c_2) phase-plane cannot intersect and exceed the curve $c_2 = w_2$ (due to conservation laws discussed in Section 4.4). Thus, the solutions that closely follow the two stationary branches $c = e_0 + z$ and $c = z_0 - w$ translate to solutions that are asymptotic to the left branch $c_2 = w_2$ and the right branch $c_2 = -w_2$ in the blown-up space. To capture this behavior, we present the following particular solution of the Riccati equation.

Theorem 4.2. The zeroth-order approximation to a solution trajectory for the desingularized system (4.52) is

$$f(w_2) = -w_2 - \sqrt{2\mu}\frac{D_{-b+1}(-\phi(w_2))}{D_{-b}(-\phi(w_2))} \tag{4.69}$$

where $\mu = k_2 E_T / 2k_1$, $b = (k_{-1} + 2k_2)/(2k_2)$ and $\phi(w_2) = \sqrt{\frac{2}{\mu}} w_2$.

Proof. To compute the asymptotics, we derive the expansions of the parabolic cylinder functions from [1, Equations 19.8.1 and 19.8.2] , as:

$$\lim_{x\to\infty} D_\nu(x) = \cos(\pi\nu)\exp\left(-\frac{x^2}{4}\right)x^\nu\left(1 - \frac{\nu(\nu-1)}{2x^2} + \dots\right), \tag{4.70a}$$

$$\lim_{x\to\infty} D_\nu(-x) = \sin\left(\pi\nu + \frac{\pi}{2}\right)\exp\left(-\frac{x^2}{4}\right)x^\nu\left(1 - \frac{\nu(\nu-1)}{2x^2} + \dots\right) + \tag{4.70b}$$

$$\frac{\sqrt{2\pi}}{\Gamma(-\nu)}\exp\left(\frac{x^2}{4}\right)x^{-\nu-1}\left(1 + \frac{(\nu+1)(\nu+2)}{2x^2} + \dots\right). \tag{4.70c}$$

Thus, taking the limit for the second term of the solution (4.53) and retaining the domi-

nant parts,

$$\lim_{w_2\to\infty} f(w_2) = -\lim_{w_2\to\infty} w_2 - \sqrt{2\mu}\lim_{w_2\to\infty} \frac{\frac{\sqrt{2\pi}}{\Gamma(b-1)}\exp\left(\frac{w_2^2}{2\mu}\right)\left(\sqrt{\frac{2}{\mu}}w_2\right)^{b-2}\left(1+\frac{\mu(b-2)(b-3)}{4w_2^2}+\dots\right)}{\frac{\sqrt{2\pi}}{\Gamma(b)}\exp\left(\frac{w_2^2}{2\mu}\right)\left(\sqrt{\frac{2}{\mu}}w_2\right)^{b-1}\left(1+\frac{\mu(b-1)(b-2)}{4w_2^2}+\dots\right)}, \tag{4.71a}$$

$$= -\lim_{w_2\to\infty} w_2 - \lim_{w_2\to\infty}\frac{\mu}{w_2}\left(\frac{\Gamma(b)}{\Gamma(b-1)}\right), \tag{4.71b}$$

$$= -\lim_{w_2\to\infty} w_2 - \frac{\mu}{b-1}\lim_{w_2\to\infty}\frac{1}{w_2}, \tag{4.71c}$$

we obtain that

$$f(w_2) \to -w_2 \text{ as } w_2 \to \infty. \tag{4.72}$$

This suggests that the solution converges to the right branch of the critical set in the blown-up space as long as $\beta_2 \neq 0$.

Similarly, let us take the limit to $w_2 \to -\infty$ to verify the convergence to the left branch of the critical manifold. Note that the term $D_\mu(z)$ grows exponentially in this case. If the terms involving $D_\mu(z)$ remain in our solution, we would obtain the behavior $f(w_2) \to \infty$ as $w_2 \to -\infty$ which does not satisfy our system. Thus, we test the particular solution obtained for $\beta_1 = 0$ and $\beta_2 \neq 0$. The solution becomes

$$f(w_2) = -w_2 - \sqrt{2\mu}\frac{D_{-b+1}\left(-\sqrt{\frac{2}{\mu}}w_2\right)}{D_{-b}\left(-\sqrt{\frac{2}{\mu}}w_2\right)}. \tag{4.73}$$

Now, taking the limit $w_2 \to -\infty$,

$$\lim_{w_2\to-\infty} f(w_2) = -\lim_{w_2\to-\infty} w_2 - \lim_{w_2\to-\infty} \frac{\sqrt{2\mu}D_{-b+1}\left(-\sqrt{\frac{2}{\mu}}w_2\right)}{D_{-b}\left(-\sqrt{\frac{2}{\mu}}w_2\right)}, \tag{4.74a}$$

$$= -\lim_{w_2\to-\infty} w_2 - \lim_{w_2\to\infty} \frac{\sqrt{2\mu}D_{-b+1}\left(\sqrt{\frac{2}{\mu}}w_2\right)}{D_{-b}\left(\sqrt{\frac{2}{\mu}}w_2\right)}, \tag{4.74b}$$

$$= -\lim_{w_2\to-\infty} w_2 - \lim_{w_2\to-\infty} \frac{\sqrt{2\mu}\cos(\pi-\pi b)\exp\left(-\frac{w_2^2}{2\mu}\right)\left(\sqrt{\frac{2}{\mu}}w_2\right)^{-b+1}}{\cos(-\pi b)\exp\left(-\frac{w_2^2}{2\mu}\right)\left(\sqrt{\frac{2}{\mu}}w_2\right)^{-b}}, \tag{4.74c}$$

$$= \lim_{w_2\to-\infty}(-w_2 + 2w_2), \tag{4.74d}$$

we obtain that

$$f(w_2) \to w_2 \text{ as } w_2 \to -\infty. \tag{4.75}$$

This suggests that the solution converges to the left branch of the critical set in the blown-up space when $\beta_1 = 0$. Here, β_2 can take any value other than 0 as it doesn't show up the final solution in terms of the original coordinates. Thus, (4.69) is the particular solution of interest. □

The expression (4.69), being a zeroth-order estimate, lies within a distance of $\mathcal{O}(r_2)$ from the true solution trajectories of the system. Figure 4.5 shows numerical evidence for its accuracy as compared to the numerical solution of system (4.49) along with the blown-up reduced systems. Its primary utility lies in the quantitative insights it offers into the dynamics, particularly near the bifurcation point. As shown in Section 4.4, the solutions intersect the nullcline curve at the peak point corresponding to the maximum complex concentration, a quantity of significant experimental relevance. Using this solution, we estimate the location of the peak point in the phase plane to investigate the validity of the reverse QSSA.

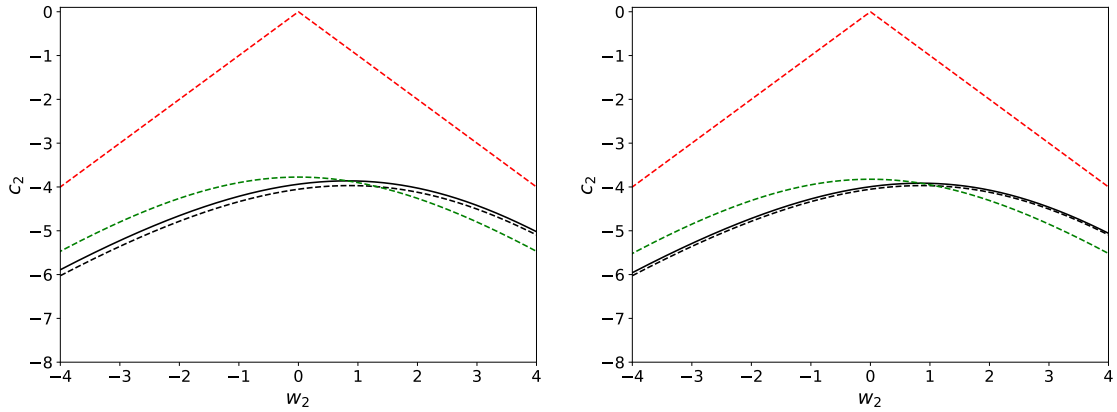


Figure 4.5: **The zeroth-order solution estimate for the desingularized blown-up system improves as r_2 (or equivalently ε) decreases.** The solid black line is the numerical solution to system (4.49) with singularity $r_2 = \sqrt{\varepsilon}$. The dashed black line is the solution to the desingularized zeroth-order system (4.69). The red dashed line is the numerical solution to the reverse QSSA blown-up to the chart K_2. The green dashed line is the numerical solution to the classical QSSA blown-up to the chart K_2. **Both**: Parameter values used are $z_0 = 10, e_0 = 5, k_1 = 1, k_{-1} = 1, k_2 = 1$, Initial Conditions used are $w(0) = -(z_0 - e_0)/(2r_2), c(0) = -E_T/(2r_2)$. **Left**: $r_2 = 0.1$. **Right**: $r_2 = 0.05$.

4.5.3 Validity of the Reverse QSS Approximation

The peak of the solution trajectories provide a way to realistically estimate the limiting (maximum) rate of peptide formation. As c^{peak} is the maximum complex concentration the systems' solutions achieve, it becomes a crucial quantity in experimental applications. Not just that, the peak value also pertains to the utility of the reverse QSSA in system approximation. Typically, the reverse QSSA is utilized for fitting experimental data and estimating the catalytic rate constant, k_2. To accurately switch between the two branches of the reverse-QSSA, we ought to have an accurate understanding of when the solutions achieve this peak:

Proposition 4.2. The peak of the solution trajectory in the blown-up chart K_2 is approximately

$$w_2^{\text{peak}} = \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \qquad c_2^{\text{peak}} = f(w_2^{\text{peak}}) \tag{4.76}$$

where $f(w_2)$ is given by (4.69). The constants are as follows

$$
\begin{aligned}
b &= \frac{k_{-1}+2k_2}{2k_2},\\
\gamma_1 &= \frac{\Gamma((b+1)/2)}{\Gamma(b/2)},\\
A &= \frac{8}{\mu}\left(\frac{b(b-1)}{2} - (4b-2)\gamma_1^2 + 6\gamma_1^4\right),\\
B &= \frac{4}{\sqrt{\mu}}\gamma_1(2b-1-4\gamma_1^2),\\
C &= -(2b-1-4\gamma_1^2).
\end{aligned}
$$

Proof. We estimate the peak to be close to the origin $(w_2, c_2) = (0, 0)$. Consider the series expansion of (4.69) near $w_2 = 0$:

$$f(w_2) \sim -w_2 - 2\sqrt{\mu}\frac{\Gamma((b+1)/2)}{\Gamma(b/2)} - 4\left(\frac{b-1}{2} - \frac{\Gamma((b+1)/2)^2}{\Gamma(b/2)^2}\right)w_2 \tag{4.78a}$$

$$- \frac{2}{\sqrt{\mu}}\left(\frac{b-1}{2}\right)\left(-4\frac{\Gamma((b+1)/2)}{\Gamma(b/2)} - \frac{\Gamma((b-1)/2)}{\Gamma(b/2)} + 4\frac{\Gamma((b-1)/2)\Gamma((b+1)/2)^2}{\Gamma(b/2)^3}\right)w_2^2 \tag{4.78b}$$

$$+ \frac{8}{3\mu}\left(\frac{b-1}{2}\right)\left(b - (b+1)\frac{\Gamma((b-1)/2)\Gamma((b+1)/2)}{\Gamma(b/2)^2} - 6\frac{\Gamma((b+1)/2)^2}{\Gamma(b/2)^2}\right. \tag{4.78c}$$

$$\left. +6\frac{\Gamma((b-1)/2)\Gamma((b+1)/2)^3}{\Gamma(b/2)^4}\right)w_2^3 + \mathcal{O}(w_2^4). \tag{4.78d}$$

Let us substitute constants $\gamma_1 = \frac{\Gamma((b+1)/2)}{\Gamma(b/2)}$ and $\gamma_2 = \frac{\Gamma((b-1)/2)}{\Gamma(b/2)}$ to simplify the above expression. Also note that the constants are related to each other as $\gamma_2 = 2\gamma_1/(b-1)$.

Rewriting the expansion, we obtain

$$f(w_2) \sim -2\sqrt{\mu}\gamma_1 - (2b-1-4\gamma_1^2)w_2 - \frac{2}{\sqrt{\mu}}\left(\frac{b-1}{2}\right)(-4\gamma_1-\gamma_2+4\gamma_2\gamma_1^2)w_2^2 \tag{4.79a}$$

$$+ \frac{8}{3\mu}\left(\frac{b-1}{2}\right)(b-(b+1)\gamma_1\gamma_2 - 6\gamma_1^2 + 6\gamma_2\gamma_1^3)w_2^3 + \mathcal{O}(w_2^4), \tag{4.79b}$$

$$\sim -2\sqrt{\mu}\gamma_1 - (2b-1-4\gamma_1^2)w_2 - \frac{2}{\sqrt{\mu}}(-2(b-1)\gamma_1 - \gamma_1 + 4\gamma_1^3)w_2^2 \tag{4.79c}$$

$$+ \frac{8}{3\mu}\left(\frac{b(b-1)}{2} - (b+1)\gamma_1^2 - 3(b-1)\gamma_1^2 + 6\gamma_1^4\right)w_2^3 + \mathcal{O}(w_2^4), \tag{4.79d}$$

$$\sim -2\sqrt{\mu}\gamma_1 - (2b-1-4\gamma_1^2)w_2 + \frac{2}{\sqrt{\mu}}\gamma_1(2b-1-4\gamma_1^2)w_2^2 \tag{4.79e}$$

$$+ \frac{8}{3\mu}\left(\frac{b(b-1)}{2} - (4b-2)\gamma_1^2 + 6\gamma_1^4\right)w_2^3 + \mathcal{O}(w_2^4). \tag{4.79f}$$

Now, differentiating the function with respect to w_2 and equating the derivative to 0, we get $Aw_2^2 + Bw_2 + C = 0$ where

$$A = \frac{8}{\mu}\left(\frac{b(b-1)}{2} - (4b-2)\gamma_1^2 + 6\gamma_1^4\right),$$
$$B = \frac{4}{\sqrt{\mu}}\gamma_1(2b-1-4\gamma_1^2),$$
$$C = -(2b-1-4\gamma_1^2).$$

Thus, the peak will occur at (4.76). □

It is worth highlighting that the functions $f(w_2)$ (4.69) and its derivative are analytic functions at $w_2 = 0$ implying the Taylor's series expansion converges to $(\mathrm{d}f/\mathrm{d}w_2)(0)$ as $w_2 \to 0$.[15] Thus, the peak estimate improves as we consider more terms from the Taylor's series when computing $\mathrm{d}f/\mathrm{d}w_2 = 0$.

To examine the accuracy of the peak estimate, we *blow-down* the values to obtain an estimate in the original coordinates, resulting in the following.

[15]The parabolic cylinder functions are analytic at $w_2 = 0$. Further, the denominator of the second term in $f(w_2)$ is not equal to 0 for $w_2 = 0$. A detailed proof on analyticity of reciprocal of functions can be found in [84].

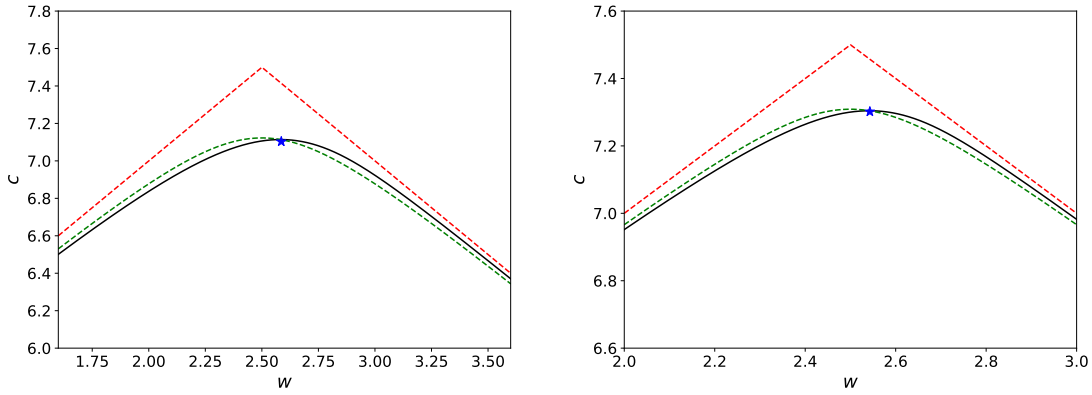


Figure 4.6: **The peak point estimate** (4.76) **for the solution trajectory improves as** ε **decreases.** The solid black line is the numerical solution to system (4.39). The red dashed line is the numerical solution to the reverse QSSA (4.44). The green dashed line is the numerical solution to the classical QSSA blown-up (4.45). **Both**: Parameter values used are $z_0 = 10, e_0 = 5, k_1 = 1, k_{-1} = 1, k_2 = 1$, Initial Conditions used are $w(0) = 0, c(0) = 0$. **Left**: $\varepsilon = 0.01$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (1.29 \times 10^{-3}, 1.51 \times 10^{-3})$; **Right**: $\varepsilon = 0.0025$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (3.07 \times 10^{-4}, 3.71 \times 10^{-4})$.

Proposition 4.3. The peak of the solution trajectory for system (4.39) is approximately

$$(w^{\text{peak}}, c^{\text{peak}}) \approx ((z_0 - e_0)/2 + \sqrt{\varepsilon} w_2^{\text{peak}}, E_T/2 + \sqrt{\varepsilon} c_2^{\text{peak}}) \tag{4.81}$$

where w_2^{peak} and c_2^{peak} are given by proposition 4.76.

Proof. We can apply the coordinate transformations in the reverse order (reverse of rescaling and shifting) to (4.76) to obtain the result. □

Figure 4.6 shows the peak estimate overlaid with numerical solutions for several parametric cases. As expected, the estimate is accurate for small values of ε and improves as $\varepsilon \to 0$. This leads to another key quantitative insight derived from the solution (4.69).

Theoretical results for fast/slow systems in standard form indicate that trajectories pass near a bifurcation point at a distance of order $\sqrt{\varepsilon}$, where ε is the small non-dimensional perturbation parameter [12]. Our peak estimate (4.76) is consistent with this: it shows that the solution passes near the bifurcation point at a distance of order $\sqrt{\varepsilon}$, reinforcing the

accuracy of the reverse QSSA in this region.

Moreover, the peak coordinates w_2^{peak} and c_2^{peak} scale as $\mathcal{O}(\sqrt{\mu})$, provided $k_{-1} = 0$ or $k_{-1} \sim k_2$.[16] These are the regimes where a bifurcation point exists. We do not consider the case $k_{-1} \gg k_2$ for now, as it eliminates the bifurcation entirely (see Section 4.6 for a discussion on this phenomenon). Therefore, the accuracy of the reverse QSSA also improves as $\sqrt{\mu}$ decreases.

To express this quantitatively, we scale the quantity by the maximum plausible distance between the nullcline and the bifurcation point, namely $E_T/2$. This yields the dimensionless parameter:

$$\frac{2\sqrt{\mu}}{E_T} = \sqrt{\frac{2k_2}{k_1 E_T}} = \sqrt{\frac{2K}{E_T}}. \tag{4.82}$$

This recovers and extends the candidate epsilon, $\tilde{\varepsilon} = 2K/E_T \quad (K = k_2/k_1)$, identified by Eilertsen et al. [46] in their analysis of the irreversible IAZA reaction mechanism (i.e., the case $k_{-1} = 0$), specifically to include the cases where $k_{-1} \sim k_2$. In these cases, the accuracy of the reverse QSSA near the bifurcation point improves as $\tilde{\varepsilon} \to 0$, which includes biologically relevant scenarios such as high enzyme concentrations [162].

What makes the IAZA mechanism and its reduced models particularly compelling is how dramatically the dynamics change with variations in k_{-1}. As we've just discovered, the accuracy of the reverse QSSA is highly sensitive to the relationship between k_{-1} and k_2. Given our understanding of the bifurcation structure and its dependence on the rate constants, can we further elucidate the phase-plane dynamics under different parametric regimes? Moreover, how does the classical nullcline-based reduction fit into approximating the behavior of the IAZA mechanism? We explore these questions in the next section.

4.6 Dynamic Imperfection, Faux Canard and More

Krupa and Szmolyan [103] perform a rigorous analysis of singularly perturbed dynamical systems near a transcritical bifurcation, employing the blow-up method on systems expressed

[16] $b = 1$ for $k_{-1} = 0$ and $b \sim 3/2$ for $k_{-1} \sim k_2$, see theorem 4.2.

in their canonical form. To relate their results to our context, we first derive a normal form for the IAZA mechanism.

For the shifted system (4.46), we introduce the new variables $\widehat{x} = k_1\widehat{c}$, $\widehat{y} = -k_1\widehat{w}$ and $\widehat{\varepsilon} = -(k_1k_2E_T/2)\varepsilon$ which yield the transformed system,

$$\dot{\widehat{x}} = k_1\left(k_1\left(\frac{\widehat{x}^2}{k_1^2} - \frac{\widehat{y}^2}{k_1^2}\right) - \left(-\frac{2\widehat{\varepsilon}}{k_1k_2E_T}\right)(k_{-1}+k_2)\left(\frac{\widehat{x}}{k_1} + \frac{E_T}{2}\right)\right), \tag{4.83a}$$

$$= \widehat{x}^2 - \widehat{y}^2 + \frac{K_M}{K}\widehat{\varepsilon} + h_1(\widehat{x},\widehat{y},\widehat{\varepsilon}), \tag{4.83b}$$

$$\dot{\widehat{y}} = -k_1k_2\left(-\frac{2\widehat{\varepsilon}}{k_1k_2E_T}\right)\left(\frac{\widehat{x}}{k_1} + \frac{E_T}{2}\right), \tag{4.83c}$$

$$= \widehat{\varepsilon}(1 + h_2(\widehat{x},\widehat{y},\widehat{\varepsilon})), \tag{4.83d}$$

where h_1 and h_2 represent higher-order perturbation terms. The system in $(\widehat{x},\widehat{y})$ is now in normal form, refer to [103, 106].

The natural question arises: do the normal form results extend directly to the IAZA mechanism, especially given the sensitivity of its dynamics to variations in k_{-1}? Recall that the critical manifold consists of two attracting and two repelling branches (see Figure 4.3). For small $\varepsilon > 0$, one might expect the slow manifold to closely track the attracting branches. However, Krupa and Szmolyan [103] show that this behavior only holds under specific parametric constraints in the normal form.

For the IAZA mechanism, these constraints translate to $K_M < K$, a condition that is not biologically realistic. In fact, for a biologically relevant regime where $K_M > K$, their analysis suggests that solutions may be repelled away from the attracting branches. Yet, such behavior is not observed experimentally, prompting the question: what is causing this deviation?

One plausible explanation is that increasing k_{-1} introduces a *dynamic imperfection* in the system. As k_{-1} grows, the transcritical bifurcation is weakened or entirely removed, fundamentally altering the system's local structure [46], resembling the case of imperfect bifurcations [67]. This imperfection likely breaks the assumptions required for the normal

form analysis to apply directly. Understanding how such imperfections affect the applicability of singular perturbation results is critical for accurately interpreting the dynamics and for justifying reduced models like the classical QSSA. In the following subsections, we analyze two cases of varying k_{-1} while k_2 remains small: $k_{-1} = 0$ and $k_{-1} \gg k_2$.

4.6.1 The Irreversible IAZA Mechanism

The Irreversible IAZA mechanism corresponds to the case when no dissociation of the complex is allowed i.e., when $k_{-1} = 0$. This is analogous to the Van Slyke–Cullen reaction mechanism and corresponds to the perturbed system:

$$\dot{w} = \varepsilon k_2 c, \tag{4.84a}$$

$$\dot{c} = k_1(e_0 + w - c)(z_0 - w - c) - \varepsilon k_2 c, \tag{4.84b}$$

which still admits a transcritical bifurcation point at $(w^*, c^*) = ((z_0 - e_0)/2, E_T/2)$ for $k_2 \to \varepsilon k_2$. Consequently, the blow-up results from Section 4.5.2 still apply to the irreversible mechanism.

From Theorem 4.2, we obtain an analytical estimate for the solution trajectory in the blown-up rescaling chart:

$$\gamma(w_2) := c_2 = -w_2 - \left(\frac{2}{\sqrt{\pi}}\right)\left(\frac{\exp(-\phi(w_2)^2)}{1 + \operatorname{erf}(\phi(w_2))}\right)\sqrt{\mu}, \tag{4.85}$$

since $b = (k_{-1} + 2k_2)/(2k_2) = 1$.[17] This also yields an estimate for the peak of the solution trajectory in the irreversible case:

$$(w^{peak}, c^{peak}) \approx ((z_0 - e_0)/2 + 0.387281\sqrt{\varepsilon\mu}, E_T/2 - 1.07312\sqrt{\varepsilon\mu}), \tag{4.86}$$

numerical accuracy of which is shown in Figure 4.7. The asymptotic accuracy of the reverse QSSA for the irreversible mechanism is limited by a term of $\mathcal{O}(\sqrt{\varepsilon})$ where ε is given by

[17] This is derived from $D_0(-x) = \exp(-x^2/4)$ and $D_{-1}(-x) = \sqrt{\pi/2}\exp(x^2/4)(1 + \operatorname{erf}(x^2/2))$.

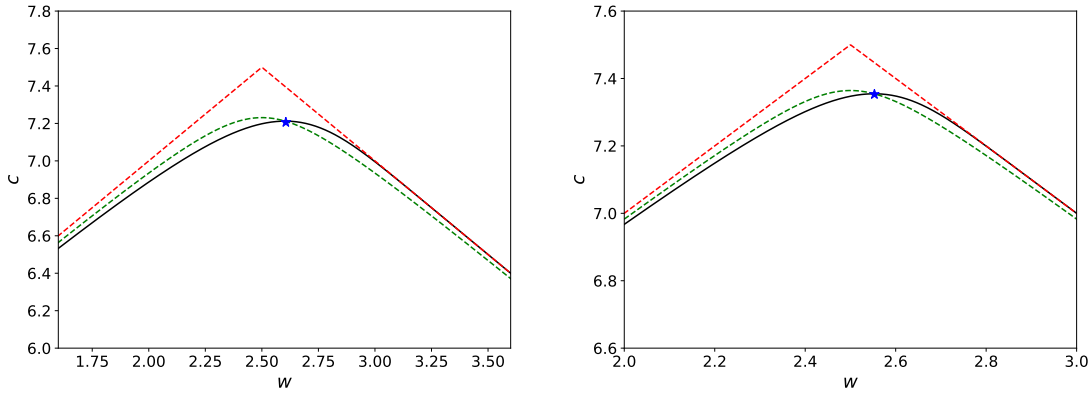


Figure 4.7: **The peak point estimate** (4.86) **for the solution trajectory for irreversible IAZA mechanism improves as ε decreases.** The solid black line is the numerical solution to system (4.84). The red dashed line is the numerical solution to the reverse QSSA (4.44). The green dashed line is the numerical solution to the classical QSSA blown-up (4.45). **Both**: Parameter values used are $z_0 = 10, e_0 = 5, k_1 = 1, k_{-1} = 0, k_2 = 1$, Initial Conditions used are $w(0) = 0, c(0) = 0$. **Left**: $\varepsilon = 0.01$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (2.06 \times 10^{-3}, 8.66 \times 10^{-4})$; **Right**: $\varepsilon = 0.0025$, Relative Error between (4.76) and numerical peak estimate is $(\Delta w^{\text{peak}}, \Delta c^{\text{peak}}) = (6.72 \times 10^{-5}, 2.14 \times 10^{-4})$.

(4.82). Also see [46] for a detailed analysis of the reverse and nullcline reduction for the irreversible mechanism.

More interestingly, the irreversible case of the IAZA mechanism is the only scenario in which the normal form results apply directly, specifically when the ratio $K_M/K = 1$. In their foundational work, Krupa and Szmolyan identify this as a degenerate case, that might admit the existence of canard solutions in the phase plane [103, 104]. For the IAZA mechanism, we find a faux canard solution trajectory that travels from the unstable to the stable branch, surprisingly passing directly through the bifurcation point, see Figure 4.8. This occurs because the curve $c = z_0 - w$ is invariant here and results in the following.

Theorem 4.3. For system (4.84), solution trajectories with initial conditions on the invariant curve $c = z_0 - w$ act as faux canards.

Proof. We show that the set

$$\Delta = \{(w, c) \in \mathbb{R}^2_{\geq 0} \ : \ c = z_0 - w\}$$

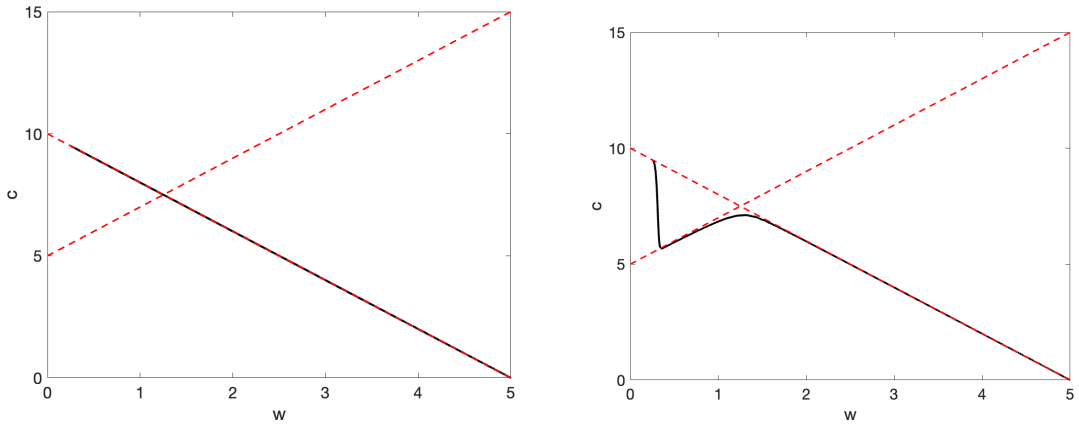


Figure 4.8: **The irreversible IAZA mechanism conforms with the normal form of a transcritical bifurcation and admits faux canard in the phase-plane.** For the same initial conditions, the reversible IAZA mechanism does not exhibit canard-like behavior. The solid black line is the numerical solution to system (4.39). The red dashed line is the numerical solution to the reverse QSSA (4.44). **Both:** The parameter values are $z_0 = 10, e_0 = 5, k_1 = 1, k_2 = 1, \varepsilon = 0.01$, The initial conditions are $w(0) = 0.5, c(0) = 9.5$. **Left:** The irreversible case corresponds to $k_{-1} = 0$. **Right:** The reversible case corresponds to $k_{-1} = 1$.

is invariant for system (4.84). For this, consider the function $V(w, c) = z_0 - w - c$ such that

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\partial V}{\partial w}\frac{\mathrm{d}w}{\mathrm{d}t} + \frac{\partial V}{\partial c}\frac{\mathrm{d}c}{\mathrm{d}t} = -\frac{\mathrm{d}w}{\mathrm{d}t} - \frac{\mathrm{d}c}{\mathrm{d}t}.$$

For $(w, c) \in \Delta$,

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -\varepsilon k_2(z_0 - w) + \varepsilon k_2(z_0 - w) = 0$$

implying that solutions with initial conditions in Δ remain in Δ for $t \to \infty$. Similar argument can be used to show invariance of Δ for $t \to -\infty$. □

In contrast, the reversible IAZA mechanism does not exhibit such behavior. All solutions get attracted to the attracting branches of the critical set (4.40) bounded away from the bifurcation point at a specific distance instead, see Figure 4.8.

This distinction complicates the construction of reduced models, particularly as we move away from the singular limit. To reiterate, the perturbative case with small k_{-1} and k_2 has been successfully approximated using the reverse QSSA and nullcline-based reductions (see

Section 4.5.2). However, if $k_{-1} \ll k_2$, the system may behave more like the irreversible mechanism and potentially exhibit canard-like behavior for some initial conditions. This introduces additional challenges for model reduction and highlights the need for further theoretical analysis.

4.6.2 Validity of the Classical Reduction

On the other hand, if $k_{-1} \gg k_2$, the system transitions into a fundamentally different singular perturbation regime governed by small k_2. In this case, the IAZA model admits the following perturbed system:

$$\dot{w} = \varepsilon k_2 c, \tag{4.87}$$

$$\dot{c} = k_1(e_0 + w - c)(z_0 - w - c) - k_{-1}c - \varepsilon k_2 c. \tag{4.88}$$

Surprisingly, this variation eliminates the transcritical bifurcation that characterizes the reversible IAZA mechanism for small k_{-1} and small k_2. As a result, the system now admits a normally hyperbolic and attracting critical manifold,

$$\Gamma_0 = \{(w, c) \in \mathbb{R}^2_{\geq 0} : c = h^-(w; K_S)\}, \tag{4.89}$$

where $h^-(w; K_S)$ is the lower branch of the quadratic nullcline obtained by replacing K_M by $K_S = k_{-1}/k_1$ in (4.21a). In this setting, the singularity is removed and there is no longer a need for blow-up analysis or desingularization. Instead, Fenichel theory applies directly, and it guarantees that the slow manifold persists and stays close to $h^-(w; K_S)$ as $\varepsilon \to 0$.

Qualitatively, as k_{-1} becomes significantly larger than k_2, we observe that solution trajectories move away from the reverse QSSA curves (4.42) and tend closer to the nullcline curve (4.21a), as shown in Figure 4.9. Now that we understand the imperfection introduced by varying k_{-1}, there is in fact a simple explanation to this.

Even in the non-hyperbolic perturbation scenario given by (4.39), if $k_{-1} \gg k_2$ i.e., the

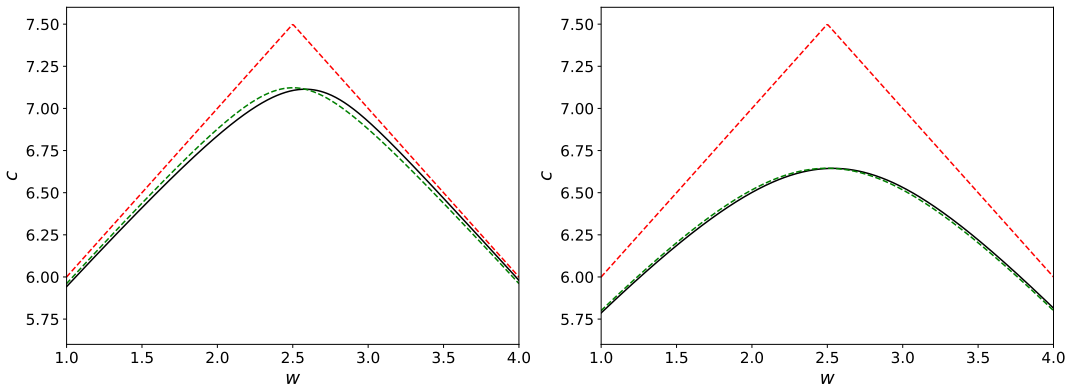


Figure 4.9: **As k_{-1} increases, the classical nullcline-based reduction provides more accurate approximation as compared to the reverse QSSA for the IAZA mechanism.** The solid black line is the numerical solution to system (4.39). The red dashed line is the numerical solution to the reverse QSSA (4.44). The green dashed line is the numerical solution to the classical QSSA blown-up (4.45). **Both:** The parameter values are $z_0 = 10, e_0 = 5, k_1 = 1, k_2 = 1, \varepsilon = 0.01$, The initial conditions are $w(0) = 0, c(0) = 0$. **Left:** $k_{-1} = 1$. **Right:** $k_{-1} = 10$. Both panels are plotted on the same axis range and scale to highlight the difference in approximation accuracy.

two parameters no longer share the same scale of "smallness", the system behaves more like the normally hyperbolic regime described by (4.87). In this limit, it becomes possible to choose a suitably small ε such that the curve $h^-(w; K_S)$ lies arbitrarily close to the classical reduction curve $h^-(w)$ due to continuity in k_2.

This implies that in the overlapping regime of small k_{-1} and small k_2[18] but $k_{-1} \gg k_2$ the classical nullcline-based reduction becomes highly accurate. Hence, the choice of an appropriate reduced model depends not only on the smallness of ε, but critically on the relative scaling of k_{-1} and k_2 as well. Understanding this scaling relationship is key to selecting the right approximation framework and avoiding misleading or inaccurate simplifications. The proximal existence of multiple reduced systems further ties back to our discussion of defining the validity of a reduced model as both its asymptotic accuracy and its predominance in the phase-plane, see Chapter 3.

[18]This is defined as compared to other parameters including z_0, k_1 and e_0.

4.7 Discussion

In this chapter, we present a rigorous analysis of the IAZA mechanism, focusing on the loss of normal hyperbolicity and its implications for model reduction. Extending reduction theory to dynamical systems with non-hyperbolic singularities is essential due to the prevalence of such features in biochemical reaction networks.

We combined classical phase-plane analysis with geometric singular perturbation theory (GSPT) and the blow-up method to study the dynamics near a transcritical bifurcation point. Since Fenichel theory assumes normal hyperbolicity, it cannot be directly applied near bifurcation points. Nevertheless, by modifying the framework, we justify the reverse QSSA, a reduced model that captures the system's behavior near the bifurcation. Alongside, we analyze the classical nullcline-based reduction, a heuristic model often used in biochemical contexts even when its validity is not formally guaranteed.

By applying a coordinate blow-up around the bifurcation point, we derive an analytical estimate for the system trajectory in the transformed phase space. This not only illuminates the mathematical structure but also has biochemical significance. Notably, we show that the trajectory passes the bifurcation point at a distance of order $\sqrt{\varepsilon}$, where $\varepsilon \sim (2k_2)/(k_1 E_T)$ providing a precise estimate for the validity of the reverse QSSA. Furthermore, we characterize the peak point, where the complex concentration is maximized, and connect it to the limiting reaction rate which is a key experimental observable. These insights offer practical guidance for designing experiments and choosing appropriate reduced models depending on the observed dynamics.

One of the most intriguing findings is that the IAZA mechanism does not conform to the canonical behavior predicted by normal form theory for transcritical bifurcations [103]. We identify k_{-1} as a key driving factor: when k_{-1} is less than or equivalent to k_2, the system exhibits bifurcation-driven dynamics, but when $k_{-1} \gg k_2$, the bifurcation structure weakens. In this latter regime, the system regains normal hyperbolicity, and Fenichel theory becomes applicable again. This regime shift also reinstates the validity of the classical QSSA, offering

a clean asymptotic justification for its use.

Additionally, we confirm the existence of canard solutions in the irreversible version of the IAZA mechanism through its conformity with the normal form. Some solutions pass from the repelling to attracting branches of the critical manifold and are a hallmark of degenerate singular perturbation problems. Their absence in the reversible system highlights the delicate interplay between reversibility, parameter scaling, and geometric structure. These insights collectively form a cohesive narrative around the applicability and limitations of reduced models in biochemical reaction networks near non-hyperbolic singularities. By charting the relationship between the bifurcation structure and the rate constants - particularly the role of the complex disassociation parameter k_{-1} to introduce imperfection - we bridge formal mathematical analysis with experimentally meaningful behavior.

Looking ahead, similar blow-up techniques can be extended to other enzymatic systems that exhibit critical transitions or multiple timescales, broadening the applicability of this analytical framework. Additionally, as discussed in Chapter 2, reduced models derived in the deterministic regime also play a crucial role in informing stochastic approximations, which are especially valuable in computational biology and systems-level simulations. This offers another potential extension of the work in this chapter.

CHAPTER 5

Gaining Space, not Hyperbolicity: A Blow-up Approach to Understanding Quasi-Steady-State Approximations

5.1 Introduction

Tikhonov and Fenichel theories are two of the most widely studied theories for the systematic reduction of multi-scale chemical mechanisms [173, 54, 55]. These theories establish the foundation to rigorously justify the use of quasi-steady-state (QSS) reductions in real-world applications of biochemical kinetics. The theory is particularly favorable for mechanisms that inherently exhibit disparate timescales yielding the rapid change in some reactant concentrations as compared to the other chemical species in the reaction. For reference, consider a singularly perturbed system in the standard fast-slow form,

$$\dot{z} = \quad f(z, y, \varepsilon), \qquad z \in \mathbb{R}^k, \quad f \in \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^k, \tag{5.1a}$$

$$\dot{y} = \varepsilon g(z, y, \varepsilon), \qquad y \in \mathbb{R}^{n-k},\, g \in \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n-k}, \tag{5.1b}$$

equipped with the initial conditions $z(0) = z_0$ and $y(0) = y_0$, where $0 < \varepsilon \ll 1$ is a small perturbation parameter and "$\dot{}$" denotes the differentiation with respect to time, t. The fast

or the initial transient subsystem corresponds to the layer problem,

$$\dot{z} = f(z, y, 0), \tag{5.2a}$$

$$\dot{y} = 0, \tag{5.2b}$$

obtained by setting $\varepsilon = 0$ and describes the (quasi) stationary behavior of the slow variables, y. The set of equilibrium points for the layer problem,

$$\Gamma = \{(z, y) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \ f(z, y, 0) = 0\}, \tag{5.3}$$

is known as the critical manifold. Consequently, the slow subsystem,

$$0 = f(z, y, 0), \tag{5.4a}$$

$$y' = g(z, y, 0), \tag{5.4b}$$

defines the slow flow that is invariably restricted to the critical manifold. This differential-algebraic system of equations yields the quasi-steady-state approximation over the slow timescale, $\tau = \varepsilon t$, where " $'$" denotes the differentiation with respect to τ.

Naturally, solutions rapidly evolve along the fast directions described by the layer problem until they achieve the slow flow defined by the quasi-steady state (QSS) condition. For systems in standard form (5.1), the QSS manifold, given by the algebraic constraint $f(z, y, 0) = 0$, coincides with the critical manifold and serves as the phase space on which the reduced (slow) dynamics unfold. We refer to this standard form as the Tikhonov standard form, defined by the popular theorem coined by Tikhonov [173] and later instated by Verhulst [180]. The theorem dictates that the solution of the problem (5.1) in the slow timescale τ converges to the solution of (5.4) over a finite time interval $\tau \in [\tau_0, \tau_1]$ as $\varepsilon \to 0$. Later, work by Hoppensteadt extended the theory to infinite time intervals [79]. In simpler words, the convergence of the slow flow to the QSS reduction is ensured when the perturbation

system is present in the standard form.

However, for systems not in Tikhonov standard form, the fast-slow structure is way less transparent. Generally, a differential equation system describing a chemical mechanism is instead of the form

$$\dot{x} = h(x, \varepsilon), \quad h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \tag{5.5}$$

where ε represents the scaling constant of some biochemical parameters involved. Here, a clear distinction between the fast and slow subsystems does not exist. Further, the QSS manifold may not coincide with the critical manifold, and the dynamics can no longer be fully described by classical Tikhonov-type reduction. In such cases, an important question arises: What governs the convergence of trajectories to a reduced model if the system does not exhibit a standard form structure?

This question was first addressed by Fenichel [55], Duchêne and Rouchon [36] and more recently by Noethen and Walcher [127]. Their analysis relies on the existence of a transformation such that the solutions of a non-standard system (5.5) are mapped to those of a standard system (5.1). Then, the convergence of solutions to (5.5) to the solutions of the reduced system obtained by Fenichel theory is established. However, there are some caveats.

First, the convergence is only known for initial conditions that lie in a suitable neighborhood of the critical manifold Γ. Thus, unless the initial conditions align with the critical manifold, how do we explicate the convergence of the solutions to the reduced system for problems in non-standard form? Second, the extension of Hoppensteadt's results to infinite time intervals also requires further thought specifically for non-standard systems [79]. One straight-forward way to tackle these conundrums is to determine the transformation of systems to Tikhonov's standard form. However, as Noethen and Walcher [127] point out, finding an explicit transformation can be computationally cumbersome for most systems, posing additional non-trivial problems.

Thus, in this chapter, we employ a methodical approach to investigate the convergence of solutions to their quasi-steady-state approximations. Significant real-world mechanisms,

not present in the standard form, exhibit behavior characteristic of standard-form fast-slow systems. The concentration of the fast-reacting species changes rapidly during an initial transient layer, while the slow-reacting species remains nearly constant. As the associated small parameters decrease (or equivalently as ε approaches 0), this separation in fast-slow variables becomes more pronounced. Two important enzymatic mechanisms that exhibit this interesting phenomenon are the Intermolecular Autocatalytic Zymogen Activation (IAZA) mechanism and the Michaelis–Menten (MM) mechanism.

To understand the dynamics on and away from the critical manifold, we attempt at gaining space around the critical manifold to magnify the neighboring dynamics. To this end, we use the blow up method on the critical manifolds which allows us to transform the coordinate space and use standard analysis tools to obtain a global picture of the trajectories. The blow-up method has been successfully used to analyze systems to gain hyperbolicity at non-hyperbolic singularities [103, 104]. Also, see Chapter 4 for a brief overview of the blow-up method and its applications.

In this work, we extend the applicability of the blow-up method to gain space instead of hyperbolicity in the phase-plane of the IAZA and MM mechanisms. This allows us to rigorously connect their fast and slow dynamics and justify the quasi-steady-state convergence. The overall aim is also to establish a geometrical approach to analyze non-standard perturbation systems, without adhering to non-dimensionalization or complex transformations.

5.2 Intermolecular Autocatalytic Zymogen Activation Kinetics for Small k_1

We start with recounting the IAZA dynamics in the (z, c) coordinates and the QSS reduction obtained for small k_1. We then employ the blow-up method to the critical manifold to provide a complete description of the layer problem.

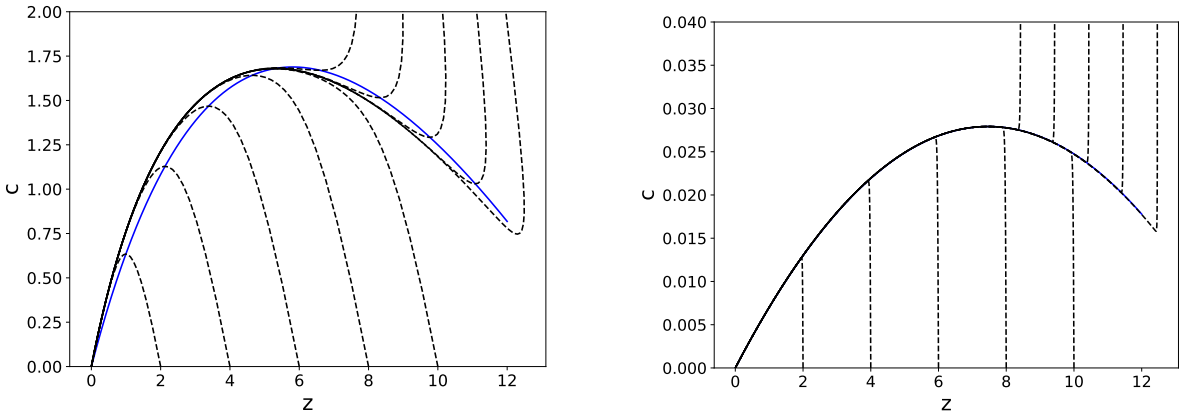


Figure 5.1: **The IAZA mechanism behaves remarkably like a system in standard form when k_1 is small.** The dashed black curves are the numerical solutions to the full IAZA ODE model (5.7) for different initial conditions. The solid blue curve is the QSS curve for the standard Fenichel reduction (5.10) for a fixed $E_T = 15$. **Left**: The parameter values are $k_1 = 0.1, k_{-1} = 1, k_2 = 1$. **Right**: The parameter values are $k_1 = 0.001, k_{-1} = 1, k_2 = 1$. **Both**: The initial conditions are $z_0 \in \{2, 4, 6, 8, 10\}$, $c_0 \in \{0, e_0\}$ and $e_0 = E_T - z_0$ such that $E_T = 15$.

5.2.1 Phase-plane Dynamics and Quasi-Steady-State Reduction

The Intermolecular Autocatalytic Zymogen Activation (IAZA) mechanism,

$$\mathrm{Z} + \mathrm{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_2} 2\,\mathrm{E} + \mathrm{W} \tag{5.6}$$

where Z is a zymogen, E is an active enzyme, C is an intermediary complex and W is a peptide, can be simply modeled using a two-dimensional system of ordinary differential equations (ODEs):

$$\dot{z} = -k_1(E_T - z)z + (k_{-1} + 2k_1 z)c, \tag{5.7a}$$

$$\dot{c} = \quad k_1(E_T - z)z - (k_{-1} + k_2 + 2k_1 z)c, \tag{5.7b}$$

where E_T is the total initial concentration of zymogen and enzyme, $E_T = z_0 + e_0$.

When k_1 is small, we employ the scaling $k_1 \to \varepsilon k_1$ where $\varepsilon \ll 1$ is a small perturbation

parameter. The mass action equations transform into a perturbation problem

$$\dot{z} = \quad k_{-1}c + \varepsilon(-k_1(E_T - z)z + 2k_1zc), \tag{5.8a}$$

$$\dot{c} = -(k_{-1} + k_2)c + \varepsilon(k_1(E_T - z)z - 2k_1zc), \tag{5.8b}$$

with the critical manifold

$$\Gamma = \{(z, c) \in \mathbb{R}^2_{\geq 0} : c = 0\}. \tag{5.9}$$

The critical manifold clearly varies from the QSS approximation curve obtained from the Fenichel's theory

$$c = \frac{z(E_T - z)}{(K_M + 2z)}, \tag{5.10}$$

where $K_M = (k_{-1} + k_2)/k_1$ is known as the Michaelis constant [118]. The corresponding *standard* quasi-steady-state approximation is,

$$\dot{z} = -\frac{k_2 z(E_T - z)}{K_M}, \tag{5.11}$$

also detailed in Eilertsen et al. [46]. Even though (5.8) doesn't exhibit the standard form structure, we observe (almost) stationary behavior of zymogen concentration as the solutions converge to the QSS curve. Figure 5.1 depicts the trajectories that get attracted to the QSS curve for small k_1.

5.2.2 Blowing-up the Zymogen Axis

As observed, the quasi-steady-state curve (5.11) might significantly differ from the critical manifold (5.9) for a system that is not present in Tikhonov standard form. Here, how do we connect the slow dynamics with the fast dynamics for a system in non-standard form? To answer this, we present a new application of the blow-up method by transforming the critical manifold for small k_1 (5.9), i.e., the zymogen coordinate axis. Using this, we study the magnified dynamics near and away from the critical manifold. Note that all the points

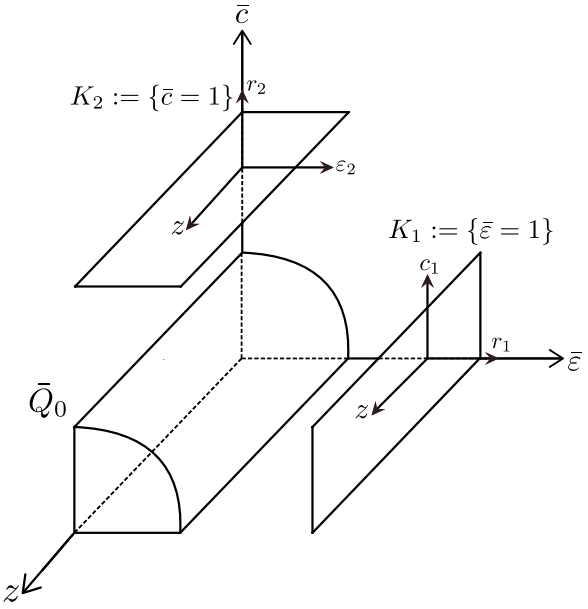


Figure 5.2: **The blown-up space can be divided into 2 coordinate charts.** The blown-up space is defined for coordinates $(z, \bar{c}, \bar{\varepsilon}, r)$ and the cylinder corresponds to the space $\mathrm{S}^1 \times \mathbb{R} \times [0, r_0]$. Chart K_1 is an intermediary blow-up in the direction of $\bar{\varepsilon}$. Chart K_2 is an intermediary blow-up in the direction of $\bar{c}$. The origin in (w, c, ε) is mapped to $\bar{Q}_0 := \mathrm{S}^1 \times \mathbb{R} \times \{r = 0\}$.

on the critical manifold are normally hyperbolic and we don't need to gain hyperbolicity with the application of the blow-up method. This is different from the blow-up method's motivation set up in Chapter 4 - our objective with the blow-up here is instead to gain 'space' in the phase-plane in order to zoom in to the local dynamics and gain insight for the global dynamics.

We recall the perturbed dynamical system for small k_1 (5.8) and write an extended and dynamically identical version including ε:

$$\begin{pmatrix} \dot{z} \\ \dot{c} \\ \dot{\varepsilon} \end{pmatrix} = \begin{pmatrix} k_{-1}c \\ -(k_{-1} + k_2)c \\ 0 \end{pmatrix} + \varepsilon \begin{pmatrix} -k_1(E_T - z)z + 2k_1zc \\ k_1(E_T - z)z - 2k_1zc \\ 0 \end{pmatrix}. \tag{5.12}$$

To blow-up the dynamics near the critical manifold, intuitively, we want to retain the

z-coordinate and transform the other coordinates around it. Thus, we utilize a cylindrical transformation which maps the z-axis in the original coordinates to a cylinder in the blown-up coordinates (and vice-versa). Consider the blow-up variables $(z, \bar{c}, \bar{\varepsilon}, r)$ which satisfy

$$\bar{c}^2 + \bar{\varepsilon}^2 = 1, \qquad r \in \mathbb{R}, \qquad z \in \mathbb{R}, \tag{5.13}$$

and achieve our aim. Thus, the map $\varphi : \mathrm{S}^1 \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^3$ defined as

$$z \to z, \qquad c \to r\bar{c}, \qquad \varepsilon \to r\bar{\varepsilon}, \tag{5.14}$$

maps the cylinder $\bar{Q}_0 := \mathrm{S}^1 \times \mathbb{R} \times \{r = 0\}$ to the z-axis, see Figure 5.2. In the blown-up space, the critical manifold will be equivalent to the conditions when $r = 0$.

As mentioned in Chapter 4, an efficient strategy is to identify coordinate charts that cover the relevant parts of the blown-up space and study the dynamics in the individual charts. There exist multiple ways to define the charts and a bit of creativity is needed to choose the best one. Since the z-coordinate will remain fixed in any chart we define, we achieve this by introducing intermediary directional blow-ups in the directions of $\bar{c}$ and $\bar{\varepsilon}$, see Figure 5.2 for a visual understanding of the coordinate charts. Also refer to Chapter 4 for an overview of directional blowups.

We obtain the first coordinate chart,

$$K_1 := \qquad z = z, \qquad c = r_1 c_1, \qquad \varepsilon = r_1, \tag{5.15}$$

in the direction of $\bar{\varepsilon}$ and the second coordinate chart in the direction of $\bar{c}$,

$$K_2 := \qquad z = z, \qquad c = r_2, \qquad \varepsilon = r_2 \varepsilon_2. \tag{5.16}$$

The steps to obtain similar coordinate charts are detailed in Appendix B. Note that the first

chart is also popularly known as a rescaling chart, as the transformations resemble a scaling transformation for the original variables.

To study the dynamics in chart K_1, consider system (5.12) with the rescaling transformation (5.15),

$$\dot{z} = -r_1 k_1 (E_T - z)z + 2r_1^2 k_1 z c_1 + k_{-1} r_1 c_1, \tag{5.17a}$$

$$\dot{c_1} = \quad k_1 (E_T - z)z - 2r_1 k_1 z c_1 - (k_{-1} + k_2)c_1, \tag{5.17b}$$

$$\dot{r_1} = \quad 0 \tag{5.17c}$$

where $r_1 = \varepsilon$ is constant. On the blown-up critical manifold (associated with the plane $r_1 = 0$), we obtain a singularly perturbed problem in standard form[1] with critical manifold

$$c_1 = \frac{z(E_T - z)}{K_M}. \tag{5.18}$$

Thus, Tikhonov's theorem [174] dictates that the approximate slow flow for $r_1 > 0$ is given by

$$\dot{z} = -r_1 . \frac{k_2 z(E_T - z)}{K_M}. \tag{5.19}$$

Moreover, we see that a degenerate scaling, $c \to r_1 c_1 \sim \varepsilon c_1$, corresponds to the dynamics in the scaling chart, K_1. This explains the distinct timescale separation between the coordinates z and c and the observed "standard form" behavior of (5.7) for small k_1; see Figure 5.1. As k_1 gets smaller, the system closely resembles a system in standard form in the sense that z is nearly - or at least appears to be nearly - constant in the approach to the QSS curve. This is also similar to the transformation idea outlined by Noethen and Walcher [127]. However, it must be emphasized that this is only a local picture of the dynamics near the critical manifold when $0 < \varepsilon \ll 1$. The global quantitative understanding of the perturbation parameter, ε, is an exhaustive field of research, briefly referred to in Chapters

[1]Here, think of r_1 as the small perturbation parameter.

3 and 4.

In chart K_2, the system translates to

$$\dot{z} = -r_2\varepsilon_2 k_1(E_T - z)z + 2r_2^2\varepsilon_2 k_1 z + k_{-1}r_2, \tag{5.20a}$$

$$\dot{r_2} = \quad r_2\varepsilon_2 k_1(E_T - z)z - 2r_2^2\varepsilon_2 k_1 z - (k_{-1} + k_2)r_2, \tag{5.20b}$$

$$\dot{\varepsilon_2} = -\varepsilon_2^2 k_1(E_T - z)z + 2r_2\varepsilon_2^2 k_1 z + (k_{-1} + k_2)\varepsilon_2 \tag{5.20c}$$

which exhibits 2 invariant subspaces $\{r_2 = 0\}$ and $\{\varepsilon_2 = 0\}$. Off the cylinder, when $r_2 \neq 0$ but $\varepsilon_2 = 0$, we observe

$$\frac{\mathrm{d}r_2}{\mathrm{d}z} = -K_M \tag{5.21}$$

which corresponds to the first integral of the layer problem for system (5.12), recall $\mathrm{d}c/\mathrm{d}z = -K_M$ for $\varepsilon = 0$. On the other hand, when $r_2 = 0$ and $\varepsilon_2 \neq 0$, we have

$$\dot{\varepsilon_2} = -\varepsilon_2^2 k_1(E_T - z)z + (k_{-1} + k_2)\varepsilon_2, \tag{5.22}$$

with the critical manifold

$$\varepsilon_2 = \frac{K_M}{z(E_T - z)}. \tag{5.23}$$

Remark 5.1. Note that the critical manifold here is the same as (5.18), just expressed in different coordinates, with $\varepsilon_2 = c_1^{-1}$ acting as a transition map between chart K_1 and K_2.[2]

Again, this corresponds to a singularly perturbed system in standard form for the coordinates (z_2, ε_2) with r_2 as the small perturbation parameter. The reduced equation becomes

$$\dot{z} = -r_2 k_2 + 2r_2^2 \frac{(k_{-1} + k_2)}{E_T - z} \tag{5.24}$$

At first glance, the reduced equation appears different from (5.19). However, substituting the transformation $r_2 = c$ recovers the conservation law $\dot{z} + \dot{w} + \dot{c} = 0$ for this system as

[2]A transition map defines the relationships of coordinates between different coordinate charts.

well, suggesting that (5.24) is a transformed version of (5.19).

Finally, the intersection of the two invariant subspaces $\{r_2 = 0\}$ and $\{\varepsilon_2 = 0\}$: $l := \{z, 0, 0\}$ is a line of equilibrium points along which the Jacobian is

$$J(z,0,0) = \begin{bmatrix} 0 & k_{-1} & 0 \\ 0 & -(k_{-1}+k_2) & 0 \\ 0 & 0 & k_{-1}+k_2 \end{bmatrix}. \tag{5.25}$$

The critical line is of saddle type such that it attracts trajectories away from the cylinder and repels trajectories that are on the cylinder towards the critical manifold. Importantly, the critical manifolds in each of the blown-up charts correspond exactly to the QSS curves of the reduced systems. Our detailed blow-up analysis shows that the system locally mimics a system in Tikhonov standard form in the vicinity of the critical manifold, particularly when the parameter k_1 is sufficiently small.

On the invariant cylinder (blown up from the critical manifold), the first integral $z = z_0$ from (5.19) governs the dynamics, reflecting the conserved quantity. Off the cylinder, the first integral obtained in (5.21) remains valid and governs the flow in the fast directions. The consistency of these first integrals across charts provides a unified description of the global phase portrait near the singularity.

These findings explain why solutions starting either on or away from the critical manifold eventually converge to the same QSS curve (5.10): the dynamics steer trajectories toward the slow manifold regardless of their initial separation from it. The blowup method provides a complete description of the layer problem that emerges when $\varepsilon = 0$ and rigorously connects the slow dynamics with the fast dynamics.

5.3 Michaelis–Menten Kinetics for Small e_0

We perform a similar analysis of the Michaelis–Menten mechanism for the perturbation scenario of small e_0. The non-standard system exhibits behavior resembling a standard form

system and a cylindrical blow-up provides better understanding of the phenomena. We focus on the calculations in this section as most of the qualitative conclusions are similar.

5.3.1 Phase-plane Dynamics and Quasi-Steady-State Reduction

The Michaelis–Menten mechanism is represented by

$$\mathrm{S} + \mathrm{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_2} \mathrm{E} + \mathrm{P} \tag{5.26}$$

where S is a substrate, E is an enzyme, C acts as an intermediary complex and P is the product and modeled using the differential equations,

$$\dot{s} = -k_1(e_0 - c)s + k_{-1}c, \tag{5.27a}$$

$$\dot{c} = \quad k_1(e_0 - c)s - (k_{-1} + k_2)c, \tag{5.27b}$$

where e_0 is the initial enzyme concentration.

When e_0 is small, we use the scaling $e_0 \to \varepsilon e_0$ to obtain the perturbed system,

$$\dot{s} = \quad (k_{-1} + k_1 s)c - \varepsilon k_1 e_0 s, \tag{5.28a}$$

$$\dot{c} = -(k_{-1} + k_2 + k_1 s)c + \varepsilon k_1 e_0 s, \tag{5.28b}$$

with the critical manifold

$$\Gamma = \{(s, c) \in \mathbb{R}^2_{\geq 0} : c = 0\}. \tag{5.29}$$

Here, the system does not admit disparate timescales in the s and c coordinates for small e_0. However, it mimics near standard form behavior for the widely-used *standard* quasi-steady-state reduction for small e_0,

$$\dot{s} = -\frac{k_2 e_0 s}{K_M + s}. \tag{5.30}$$

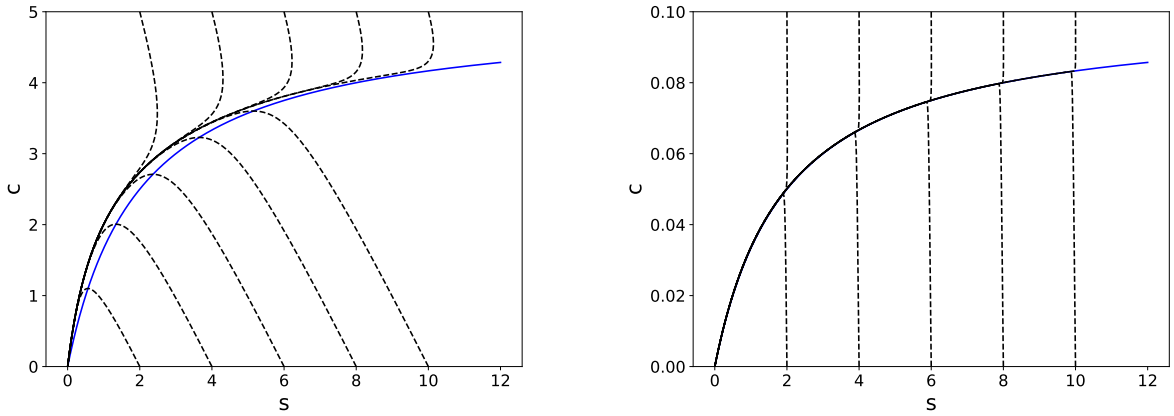


Figure 5.3: **The MM mechanism behaves remarkably like a system in standard form when e_0 is small.** The dashed black curves are the numerical solutions to the full MM ODE model (5.27) for different initial conditions. The solid blue curve is the QSS curve for the standard Fenichel reduction (5.31). **Left**: $e_0 = 5$. **Right**: $e_0 = 0.1$. **Both**: The initial conditions are $s_0 \in \{2, 4, 6, 8, 19\}$ and $c_0 \in \{0, e_0\}$. The parameter values are $k_1 = 1, k_{-1} = 1, k_2 = 1$.

Figure 5.3 shows that the trajectories both on and away from the critical manifold converge to the QSS curve,

$$c = \frac{e_0 s}{K_M + s}. \tag{5.31}$$

Owing to the wide popularity of the Michaelis–Menten system, there exist past works on the convergence of solutions to the QSS approximation. Segel and Slemrod [163] use non-dimensional scaling variables in order to heuristically define the fast and slow timescales. While effective, scaling analysis often relies on creativity and intuition, which can result in varied and occasionally inaccurate conclusions. To this end, we implement our approach of blowing-up the critical manifold to explicate the global dynamics.

5.3.2 Blowing-up the Substrate Axis

To begin with our analysis, we recall system (5.28) with appended ε,

$$\begin{pmatrix} \dot{s} \\ \dot{c} \\ \dot{\varepsilon} \end{pmatrix} = \begin{pmatrix} (k_{-1} + k_1 s)c \\ -(k_{-1} + k_2 + k_1 s)c \\ 0 \end{pmatrix} + \varepsilon \begin{pmatrix} -k_1 e_0 s \\ k_1 e_0 s \\ 0 \end{pmatrix}. \tag{5.32}$$

For our purposes, the critical manifold i.e., the c-axis is blown-up to a cylinder. Thus, we use blow-up variables which satisfy

$$\bar{c}^2 + \bar{\varepsilon}^2 = 1, \qquad r \in \mathbb{R}, \qquad s \in \mathbb{R} \tag{5.33}$$

and the map defined as

$$s \to z, \qquad c \to r\bar{c}, \qquad \varepsilon \to r\bar{\varepsilon}. \tag{5.34}$$

This transformation maps the cylinder $\bar{Q}_0 := \mathrm{S}^1 \times \mathbb{R} \times \{r = 0\}$ to the s-axis. Again, we cover the blown-up space using two coordinate charts, similar to Figure 5.2. The first coordinate chart is given by

$$K_1 := \qquad s = s, \qquad c = r_1 c_1, \qquad \varepsilon = r_1, \tag{5.35}$$

in the direction of $\bar{\varepsilon}$ and the second coordinate chart is in the direction of $\bar{c}$,

$$K_2 := \qquad s = s, \qquad c = r_2, \qquad \varepsilon = r_2 \varepsilon_2. \tag{5.36}$$

Using the transformation (5.35) on system (5.32), we obtain the new vector field in chart K_1:

$$\dot{s} = \quad r_1(k_{-1} + k_1 s)c_1 - r_1 k_1 e_0 s, \tag{5.37a}$$

$$\dot{c}_1 = -(k_{-1} + k_2 + k_1 s)c_1 + k_1 e_0 s, \tag{5.37b}$$

$$\dot{r}_1 = \quad 0 \tag{5.37c}$$

with $r_1 = \varepsilon$ being a constant. Considering r_1 as the small perturbation parameter yields a system in the Tikhonov standard form with the critical manifold

$$c_1 = \frac{e_0 s}{K_M + s}. \tag{5.38}$$

Thus, a simple application of Tikhonov's theorem provides the estimate of the slow flow for $r_1 > 0$,

$$\dot{s} = -r_1 \frac{k_2 e_0 s}{K_M + s} \tag{5.39}$$

resembling the standard QSS approximation (5.30). Once again, we observe that the scaling $c \to \varepsilon c_1$ leads to a timescale separation in s and c coordinates, resulting in the standard-form behavior. See Chapters 2 and 3 for an understanding of the small perturbation parameter ε. This presents a local picture near the critical manifold.

For chart K_2, we have the vector field

$$\dot{s} = \quad r_2(k_{-1} + k_1 s) - r_2 \varepsilon_2 k_1 e_0 s, \tag{5.40a}$$

$$\dot{r}_2 = -r_2(k_{-1} + k_2 + k_1 s) + r_2 \varepsilon_2 k_1 e_0 s, \tag{5.40b}$$

$$\dot{\varepsilon}_2 = \quad \varepsilon_2(k_{-1} + k_2 + k_1 s) - \varepsilon_2^2 k_1 e_0 s, \tag{5.40c}$$

with two invariant subspaces $\{r_2 = 0\}$ and $\{\varepsilon_2 = 0\}$. The cylinder corresponds to substituting $r_2 = 0$. Off the cylinder, with $r_2 \neq 0$ but $\varepsilon_2 = 0$, we obtain the differential equation

$$\frac{\mathrm{d}r_2}{\mathrm{d}s} = \frac{K_M + s}{K_S + s} \tag{5.41}$$

which is also the first integral of the layer problem of (5.28). Near the invariant subspace $\{r_2 = 0\}$, we retrieve a singular perturbation system in the standard form for the (s, ε_2) coordinate space. The critical manifold is

$$\varepsilon_2 = \frac{K_M + s}{e_0 s} \tag{5.42}$$

and the reduced equation becomes

$$\dot{s} = -k_2 r_2. \tag{5.43}$$

Similar to the IAZA case in Section 5.2, the critical manifold in ε_2 is the same as (5.38)

noting the transition map $\varepsilon_2 = c_1^{-1}$. Further, given the transformation $c = r_2$, the reduced equation also recovers the conservation law

$$\dot{s} + \dot{p} = -k_2 c + k_2 c = 0 \tag{5.44}$$

for the Michaelis–Menten system (5.27).

The last component is the analysis of the intersection of the cylinder with the invariant subspace $\{\varepsilon_2 = 0\}$. This connects the fast dynamics off the cylinder with the dynamics on the cylinder. Setting $r_2 = \varepsilon_2 = 0$ results in trivial dynamics and a line of equilibrium points. The Jacobian matrix along the line of equilibria is

$$J(s,0,0) = \begin{bmatrix} 0 & k_{-1} + k_1 s & 0 \\ 0 & -(k_{-1} + k_2 + k_1 s) & 0 \\ 0 & 0 & k_{-1} + k_2 + k_1 s \end{bmatrix}. \tag{5.45}$$

Consequently, the critical line is of saddle type: it attracts trajectories off the cylinder and repels trajectories on the cylinder toward the critical manifold. By blowing up the critical manifold we have shown that locally, near the critical manifold, the systems mimics a system in standard form when e_0 is sufficiently small. On the the cylinder, the first integral is $s = s_0$; off the cylinder the first integral is given by (5.41). Hence, the blowup method provides a more complete description of the layer problem that emerges when $\varepsilon = 0$.

5.4 Discussion

Our primary objective is to investigate the convergence of solutions to quasi-steady-state (QSS) approximations. When the perturbed system is not in the standard fast-slow form, Tikhonov's classical theorem on model reduction and convergence does not directly apply [173]. This limitation necessitates a deeper exploration of the underlying fast and slow dynamics, and how they interact in such systems.

In this chapter, we extend the blow-up method to hyperbolic systems to analyze the fast

Algorithm 5.1 Cylindrical Blow-up Method for 2D Models

Let (x, y, ε) represent coordinates in a two dimensional fast-slow system.
Identify $\{y = 0\}$ as the critical manifold for $\varepsilon = 0$.
Introduce blown-up coordinates $(x, \bar{y}, \bar{\varepsilon}, r)$ that satisfy

$$\bar{y}^2 + \bar{\varepsilon}^2 = 1, \qquad r \in \mathbb{R}, \qquad x \in \mathbb{R}.$$

Define the first coordinate chart that covers the direction $\bar{\varepsilon}$:

$$K_1 := \qquad x = x, \qquad y = r_1 y_1, \qquad \varepsilon = r_1.$$

Define the second coordinate chart that covers the direction $\bar{y}$:

$$K_2 := \qquad x = x, \qquad y = r_2, \qquad \varepsilon = r_2 \varepsilon_2.$$

Apply standard ODE analysis tools on the two charts and connect their dynamics to the original ODE system.

and slow dynamics in detail. A key observation is that, for systems not in Tikhonov standard form, the critical manifold and the QSS curve generally do not coincide. This distinction, although often overlooked, plays a crucial role in many biologically significant systems and their widely used QSS reductions. Consequently, it raises important questions about the conditions under which solutions still converge to QSS curves despite the lack of standard structure.

To address these questions, we examine two classical enzymatic mechanisms: the IAZA mechanism for small k_1, and the MM mechanism for small e_0. Using the blow-up technique, we analyze the dynamics near the corresponding critical manifolds and rigorously connect the behavior on and near these manifolds.

We show that in the vicinity of the critical manifolds, these systems do, in fact, locally admit a Tikhonov-like structure for small perturbation parameters. This explains the observed convergence to the QSS curve in the phase plane, even though the global system is not in standard form. Additionally, we demonstrate that the first integrals of the blown-up systems align with those of the original fast and slow subsystems, further validating the reduction.

Ultimately, our work addresses a fundamental question: why do trajectories, even those that originate away from the critical manifold, tend to be attracted to the QSS curve in the phase-plane, despite the absence of a standard fast-slow decomposition? Fenichel theory, as we know it today, does not provide a clear explanation for this phenomenon.

To formalize this approach, Algorithm 5.1 outlines a general cylindrical blow-up procedure for 2D systems. This framework can be applied broadly to detect potentially hidden slow-fast structures and assess the validity of QSS reductions in models beyond standard form assumptions.

CHAPTER 6

Gradient-based Optimization of Noisy ODE Models for Polyphasic Rodent Sleep-Wake Behavior

6.1 Introduction

Sleep-wake regulation has been extensively investigated through mathematical and statistical modeling approaches, offering insights into the underlying biophysical mechanisms [2, 17, 16]. In rodents, the irregular and polyphasic nature of sleep-wake transitions poses unique challenges. The dynamics are characterized by frequent, short wake bouts interspersed with non-rapid eye movement (NREM) and rapid eye movement (REM) sleep. While probabilistic models capture the observed noise in experimental data [136, 13, 131], an alternative and understudied strategy is to append stochastic components to otherwise deterministic models based on underlying dynamic mechanisms [34].

A central goal in the latter modeling effort is the estimation of meaningful parameters that govern neuronal or physiological dynamics. This estimation typically involves solving inverse problems using optimization techniques applied to domain-specific ordinary differential equation (ODE) models. Among the most widely used methods are Bayesian inference (e.g., MCMC, variational inference) and, more recently, neural network-based surrogate models and emulators [26, 178, 68]. Despite these advancements, relatively limited work has explored the use of gradient-based optimization in noisy, mechanistic ODE-based models, particularly in the context of sleep-wake regulation [90].

Even though gradient-based methods are cornerstones of optimization and modern ma-

chine learning, this gap is due to several intrinsic challenges. First, experimental data samples are often coarse-grained: sleep states are scored at discrete intervals (e.g., 2.5 seconds intervals) into Wake, REM, and NREM stages, typically derived from Electroencephalogram (EEG) and Electromyography (EMG) readings [21, 131]. Model outputs that resemble the discrete hypnograms that summarize sleep/wake behavior over time hinder efficient gradient computations. Moreover, fitting ODE-based models to such data requires a careful choice of appropriate error metrics for comparing model output to data, especially in the presence of noise.

To this end, we present an end-to-end framework for fitting noisy ODE-based models of a sleep/wake regulatory network to experimental rodent sleep data using gradient-based optimization. Specifically, rodent sleep exhibits unique features that complicate modeling. In contrast to the relatively consolidated sleep observed in humans, rodents frequently display fragmented sleep patterns with numerous short wake episodes interrupting longer REM and NREM bouts [176]. This irregularity is hard to capture through deterministic models, often requiring ad hoc noise components or heuristic thresholds for transition dynamics.

To address this, we introduce modifications to a physiologically motivated ODE model by incorporating stochasticity directly into the wake-promoting neuronal population and homeostatic sleep-wake state switching. This modification yields spiky, intermittent wake activity consistent with rodent recordings. We then use this model in an optimization framework that fits the simulated outputs to empirical data via gradient descent. Our primary technical contribution lies in employing automatic differentiation for gradient computation, enabling efficient and scalable parameter estimation even in the presence of noise. Details of this implementation are provided in Section 6.4.1.

Additionally, we propose robust strategies for model evaluation, including parameter sensitivity analyses, performance metrics aligned with sleep-state classification, and data augmentation schemes tailored for experimental data variability. While our focus is on modeling rodent sleep, the broader aim is to develop a generalizable framework for gradient-based

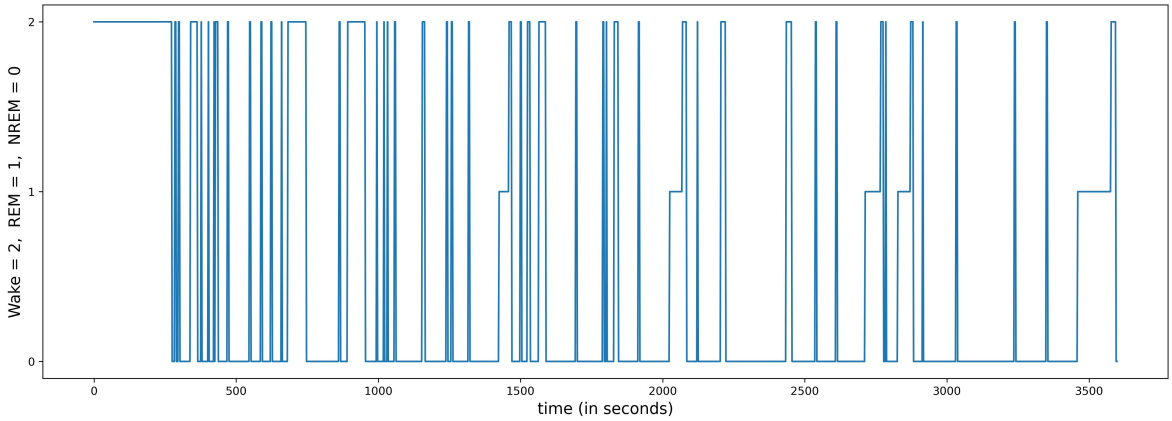


Figure 6.1: **A hypnogram depicting 1 hour of wake, REM and NREM states from sleep scored data.** Unpublished data is courtesy of the Weber Lab (U of Pennsylvania); The mouse ID is $c50_042418n1$; and the hour shown is from 7 AM to 8 AM. Mice in this dataset are exposed to a 12/12 light/dark schedule with lights turned on at 7 AM.

parameter inference in noisy ODE systems. This approach also avoids reliance on prior distributions or likelihoods, making it suitable for domains where such information is sparse or unreliable.

The rest of the chapter is structured as follows. Section 6.2 presents the experimental data used for our analysis, detailing its key characteristics and components. Section 6.3 provides an overview of the ODE model with its physiological parameters, model equations and the algorithm to append extrinsic noise. In Section 6.4, we describe all the stages of the methodology including sensitivity analysis and data processing along with the modeling and algorithm decisions for our analysis. Finally, in Section 6.5, we present a summary of our results followed by a discussion of potential improvements and future work in Section 5.4.

6.2 Experimental Data: Sleep-Wake Behavior in Rodents

In this section, we present the details of the wildtype mice data from the Weber lab (University of Pennsylvania). These data were collected to observe their sleep schedules using three states: Wake, Non-REM and REM [131]. This will help us understand the characteristics of the data that should be considered in a robust model-based analysis.

The dataset comprises 124 recordings in total, classified between the different states using

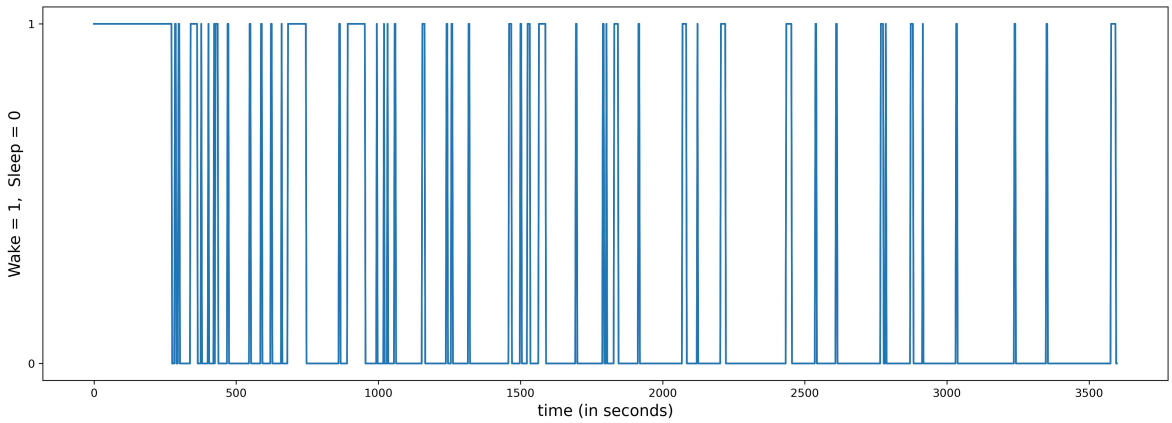


Figure 6.2: **A hypnogram depicting 1 hour of sleep/wake states by combining NREM and REM together from sleep score data.** We also relabel wake recordings with 1 and sleep recordings as 0. The mouse ID is $c50_042418n1$ and the hour shown is from 7 AM to 8 AM.

discrete labels. Here, a score of 2 stands for wake, 1 stands for NREM and 0 denotes REM.[1] The scores for the hypnograms are determined based on the observed EEG and EMG spectograms with a time resolution of 2.5 seconds, refer to Chapter 1 for more details on data retrieval. See Figure 6.1 for an example hypnogram and Park et al. [131] for more details.

A crucial component of the sleep dataset is the light cycle to which the mice are exposed to during the experiments. For our dataset, the lights were kept on for 12 hours starting at 7 AM, a period referred to as the *light phase* for the observed data, followed by a 12 hours *dark phase* starting at 7 PM. The dataset consists of hypnograms recorded across varying durations ranging from 5 hours to 24 hours and two phases of light exposure. Since wildtype mice are nocturnal animals, they exhibit more activity in the dark phase than the light phase, making the phase distinction an important focus of modeling research [121, 41, 148]. For our analysis, we separate the data into light phase and dark phase segments. Due to variation on sleep/wake behavior across the 24-h day, we further restrict our analyses to two 2 hour intervals of the dark and light phases, respectively. Before delving into the specific

[1]The original data also classifies between subsets of NREM sleep denoted by 0,3,4,5,6,7 states. We combine the NREM observations as one label i.e., 0 for simplicity.

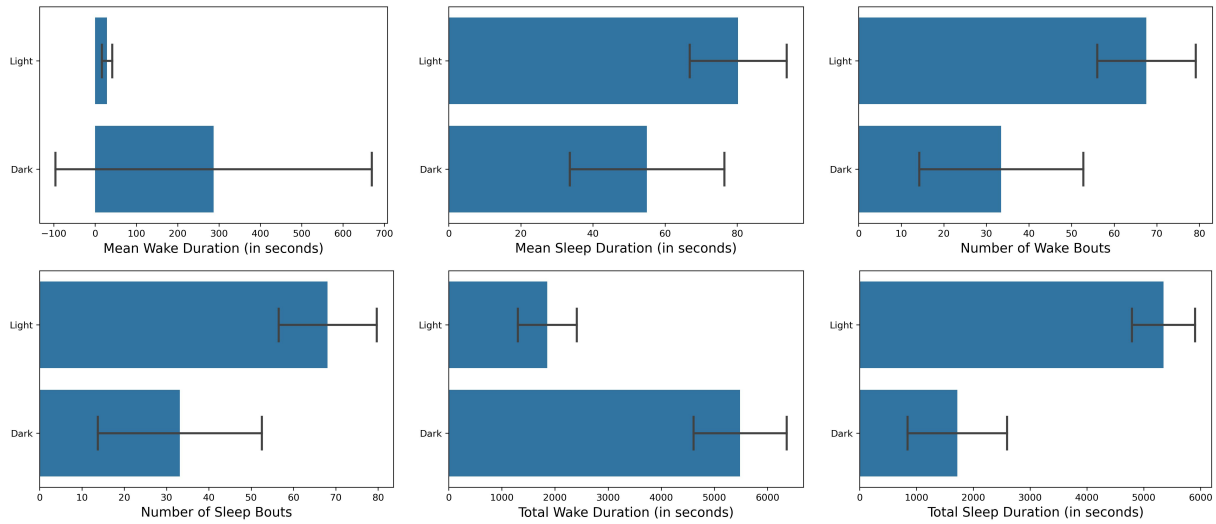


Figure 6.3: **The six inference measures vary significantly over the two observation phases.** We have removed the outlier recordings and the 6 dark phase recordings which contain only sleep. We see a negative range of mean wake durations for the dark phase possibly because there is a huge variation in the mean wake measure for the dark phase recordings.

characteristics of each phase, we note that this study adopts a simplified two-state model: sleep versus wake. Accordingly, we combine REM and NREM states into a single "sleep" category and relabel the wake state as 1 and the sleep state as 0, as illustrated in Figure 6.2.

6.2.1 Descriptive Measures of Sleep-Wake Data

To understand characteristics that describe the data, we analyze and visualize a two-hour segment from each phase. We choose 50 recordings which include at least the first four hours of the light phase and 54 recordings which include at least the first four hours of the dark phase; the two sets of recordings are disjoint.[2] However, we intentionally exclude the first two hours of each phase to avoid any initial transient behaviors that could introduce bias into our analysis; thus, we analyze the recordings from 9 AM to 11 AM for the light phase and 9 PM to 11 PM for the dark phase.

Since we adopt a two-state sleep–wake model, the durations of each state and the transitions between them are central to our study. We define a *wake bout* as a continuous period

[2]Out of the 54 dark phase recordings, 6 were completely in the sleep state from 9 PM to 11 PM. We do not include them in the forthcoming visualization.

of wake observations and a *sleep bout* as a continuous period of sleep observations. The durations of these bouts are referred to as wake and sleep durations, respectively. Each recording comprises multiple sleep and wake bouts, with some wake bouts lasting as short as 2.5 seconds. Notably, Park et al. [131] define wake bouts shorter than 20 seconds as *microarousals*, a fascinating phenomenon observed in sleep studies.

To capture the overall dynamics of the dataset, we introduce six inference measures, which will also be instrumental in evaluating our model's performance. For each two hour interval, these inference measures include the following: mean duration of wake/sleep bouts; mean number of wake/sleep bouts; and total wake/sleep duration. The mean wake duration and the mean sleep duration represent the average of the various wake/sleep bouts recorded for each mouse recording over the selected time period. The light phase corresponds to a higher number of wake bouts recorded with shorter durations, consistent with the occurrence of microarousals. In contrast, the dark phase observes a lower number of wake bouts with longer mean durations for each mouse. As expected, the total time spent in sleep is higher for the light phase as compared to the dark phase and vice versa. Figure 6.3 depicts the inference measures and compares their means and standard deviations for the two phases. Due to the stark differences in the behavior between the two phases, we expect some model parameters to change with circadian phase. Thus, we focus our subsequent analysis on the light phase, with the future goal of extending our methodology to dark phase datasets to investigate circadian variation in model parameters.

6.2.2 Underlying Distribution for Wake Activation

The presence of short-lived wake bouts in the light phase is of great interest to the field of mathematical modeling, not just due to their frequent occurrence but also because they are difficult to simulate accurately using mathematical models. Since we add exogenous noise to our model in order to emulate the irregular microarousal periods, it is crucial to understand if the wake activation events follow a specific distribution in the experimental data. This would enable us to design the noise injection more effectively and estimate the physiological

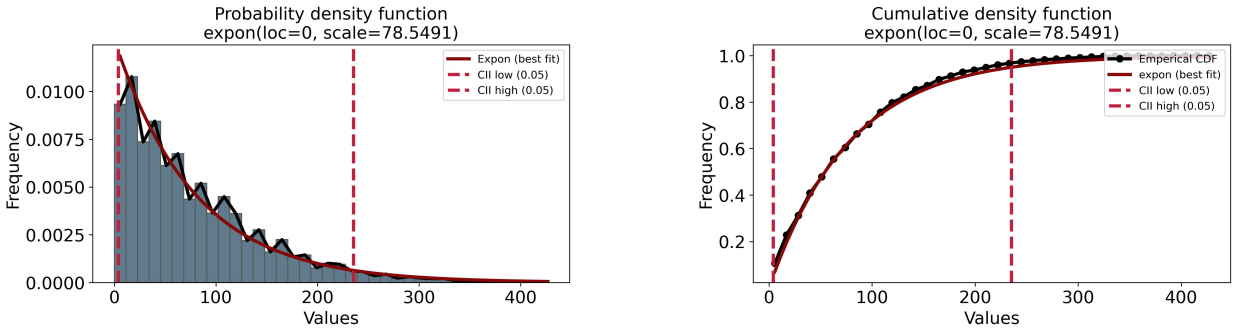


Figure 6.4: **The probability density function and the cumulative density function of the best fit theoretical exponential distribution overlap with the empirical data of sleep durations.** We use the python package 'distfit' for these plots. The x-axis denotes the values of sleep durations for the 50 light-phase data recordings from 9 AM to 11 AM. The best fit exponential distribution has the parameters loc $= 0$ and scale $= 78.549$. This corresponds to a probability density function of $\text{pdf}(x) = \exp(-(x - \text{loc}))/\text{scale}$. In terms of the rate parameter, r, this corresponds to $r = 1/\text{scale}$.

parameters accurately.

To reproduce the wake bouts or spikes shown in Figure 6.2, we apply noise stimuli to the wake neuronal population, which in turn triggers a wake activation. See Section 6.3.3 for more details. Thus, we model the noise stimuli as recurring events of a probabilistic process. Historically, neuronal spikes have been modeled using Poisson distributions with a fixed rate [5, 96], in which case the inter-event intervals follow exponential distributions.

To test this for the light phase data, we begin by extracting the inter-event intervals for the wake activation events. While it is not possible to determine the precise timing of a potential neuronal spike from discrete sleep scoring data, a straight-forward estimate is to treat the start times of wake bouts as the "probabilistic events" and the preceding sleep durations as the "inter-event intervals".

Across the 50 light-phase recordings, we identify a total of 3403 sleep bouts in the 9 AM to 11 AM window. Figure 6.4 shows the fit of the set of these durations to a theoretical exponential distribution. The data possibly contains fewer long sleep intervals than what the theoretical exponential distribution predicts. A lighter-tailed distribution, such as gamma distribution, might also describe the sleep data well.

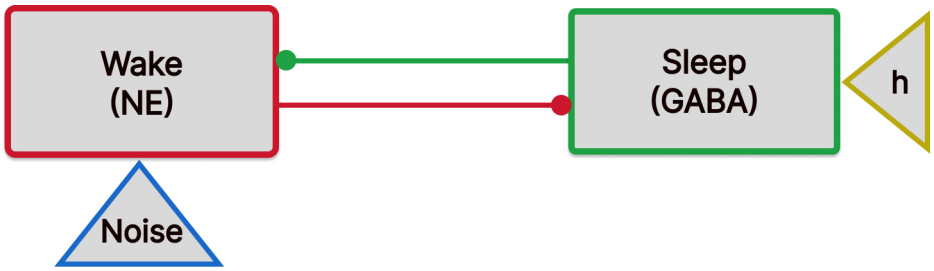


Figure 6.5: **A schematic representation for the physiological two-state model.** The wake and sleep promoting neuronal populations correspond to the neurotransmitters Noradrenaline (NE) and Gamma Aminobutyric Acid (GABA) respectively. The circles denote inhibitory synaptic connections between the two populations. h denotes the homeostatic input and Noise denotes the noise stimuli input to the wake-promoting population.

Nevertheless, since we aim to model the noise stimuli as a discrete probabilistic process, we adhere to the exponential distribution for simplicity and consistency with earlier works. See C.2 for more on this modeling choice. The skewedness doesn't impair the performance of our noise-driven model, but it remains an important reflection of the underlying biological complexity. Finally, the estimated rate of the Poisson distribution that empirically fits the current light phase data is $r = 0.0127$.

6.3 Mathematical Model: Simulating the Sleep-Wake Dynamics

In this section, we present our two-state flip-flop model of sleep/wake regulation and detail the differential equations, steady-state equations and the model parameters. We also present the different components of exogenous noise added to this ODE-based model.

6.3.1 Flip-Flop ODE Model

Based on previous works, we utilize a wake/sleep flip-flop system composed of two neural populations with one that promotes wakefulness and another that promotes sleep [34, 39]. The two populations are connected through reciprocal inhibitory pathways mediated by neurotransmitters. In the schematic in Figure 6.5, rectangular boxes denote neuronal populations with firing rates labeled as $f_W(t)$ for wake-promoting neurons and $f_S(t)$ for sleep-promoting neurons. Circular nodes indicate the concentrations of neurotransmitters released

by these neurons.

The wake-promoting regions include the locus coeruleus (LC) and the dorsal raphe (DR) whereas the sleep-promoting population represents the sleep-active neurons in the ventrolateral preoptic nucleus (VLPO) region [154, 155, 6]. These are mostly situated in the lower regions of the brain, namely the hypothalamus and the brainstem. The sleep population releases a neurotransmitter Gamma Aminobutyric Acid (GABA) with the concentration modeled as C_G, while the wake population releases the neurotransmitter Noradrenaline (NE) whose concentration is captured by the variable C_{NE}.

We model the behavior and interactions between the two neuronal populations using differential equations for their firing rates that vary over time. The neurotransmitter concentrations are driven by the firing rates of pre-synaptic (i.e., before the interaction/synapsis occurs) populations and in turn drive the firing rates of the post-synaptic neuronal populations, see Figure 1.2.

Another crucial component in the model is the homeostatic sleep drive, denoted by $h(t)$, which regulates the switch between the wake and sleep states [183, 18, 17]. Intuitively, it represents the tendency for sleep that accumulates during wakefulness and dissipates during sleep. The homeostatic input increases during the activation of the wake-promoting population. However, as $h(t)$ rises, it activates the sleep-promoting population, triggering a switch to the sleep state. Once in this sleep state, the reduced activity of the wake-promoting population leads to a decline in $h(t)$, eventually deactivating the sleep-promoting neurons and prompting a return to the wake state. Adenosine, a neuro-modulator that accumulates during prolonged wakefulness and dissipates during sleep, serves as a biological correlate of the homeostatic sleep drive and is widely used in models to represent this process [140, 141, 23, 123].

In future analysis, our model formulation for flip-flop dynamics can be extended by incorporating other biological processes such as the circadian rhythm. Notably, including the suprachiasmatic nucleus (SCN) may help elucidate the behavioral differences between the

light and dark phases. The suprachiasmatic nucleus (SCN) functions as the master circadian pacemaker, critically controlling the timing and consolidation of the sleep-wake cycle with the light-dark cycle [122, 40]. Collectively, these mechanisms have been well-studied and formalized in various mathematical models across mammalian species and shown to explain interspecies variation in sleep timing and structure [166, 146, 138].

6.3.2 Model Equations and Parameters

Formally modeling the schematic described in Figure 6.5, we denote the wake and sleep promoting neuronal populations as W and S respectively. Here, the firing rates of the neuronal populations are governed by the equations,

$$\frac{\mathrm{d}f_W}{\mathrm{d}t} = \frac{1}{\tau_W}\left(W_\infty(-g_{G,W}C_G) - f_W\right), \tag{6.1a}$$

$$\frac{\mathrm{d}f_S}{\mathrm{d}t} = \frac{1}{\tau_S}\left(S_\infty(-g_{NE,S}C_{NE}) - f_S\right), \tag{6.1b}$$

where P_∞ denotes the steady-state firing rate function

$$P_\infty(x) = \frac{P_{\max}}{2}\left(1 + \tanh\left(\frac{x-\beta_P}{\alpha_P}\right)\right). \tag{6.2}$$

Here, P is either W or S and the quantities $P_{\max}$, α_P and β_P denote the maximum firing rate, sensitivity of response and half-activation threshold, respectively [6]. The steady-state function, resembling a sigmoidal curve, has been extensively studied in firing rate modeling [137, 31, 16, 33].

In the model equations (6.1), the weight parameters $g_{n,P}$ are written such that n stands for the corresponding neurotransmitter for a given neuronal population, P. These parameters convert the neurotransmitter concentrations into synaptic input to the neuronal populations. The negative signs in front of the synaptic weights denote the mutual inhibition between the two populations.

Furthermore, we model the neurotransmitter concentrations, C_n, $n = NE,\ G$, using

another steady state function:

$$C_n = \tanh\left(\frac{f_P}{\gamma_n}\right) \tag{6.3}$$

where γ_n is the sensitivity of response for the neurotransmitter, n. Here, f_P denotes the corresponding population's firing rate i.e., $P = W$ for $n = NE$ and $P = S$ for $n = G$.

Finally, the homeostatic sleep drive is governed by the equation

$$\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{\mathcal{H}(f_W - h_{th}).(h_{max} - h)}{\tau_{hw}} + \frac{\mathcal{H}(h_{th} - f_W).(h_{min} - h)}{\tau_{hs}} \tag{6.4}$$

where $\mathcal{H}(\cdot)$ denotes a Heaviside function and h_{th} is the threshold value that determines the switch between sleep and wake. During wake, the homeostatic drive accumulates exponentially with the time constant τ_{hw} whereas during sleep, it declines exponentially with the time constant τ_{hs}. More importantly, the homeostatic drive h, modulates the activity of the sleep-promoting population by influencing the half-activation threshold,

$$\beta_S = k_2 h + k_1. \tag{6.5}$$

Note that the half-activation threshold for the wake-promoting population, β_W, is a constant.

We slightly modify the baseline parameter values from the deterministic model implemented by Athanasouli et al. [6] to better match the key characteristics of this experimental data set. For instance, we observe longer sleep bouts of around 1500 seconds instead of equal sleep and wake bout durations in the data. This is easily reflected in the ODE model by changing the parameter τ_{hs}. The complete list of parameter values used for the baseline model is presented in Table 6.1. See Figure 1.3 for a sample time trace of the deterministic model adjusted to include smaller wake cycles.

6.3.3 Addition of Noise to ODE Model

In order to simulate the variable, poly-phasic nature of rodent sleep, we add noise to the flip-flop deterministic model. In this subsection, we present the algorithmic details to

Firing Rate Parameters (6.1)			
$W_{max} = 6$ Hz	$\tau_W = 5$ seconds	$\alpha_W = 0.4$	$\beta_W = -0.3$
$S_{max} = 5$ Hz	$\tau_S = 5$ seconds	$\alpha_S = 0.2$	
Synaptic Connection Parameters (6.3)			
$g_{NE,N} = 3$	$\gamma_{NE} = 5$		
$g_{G,W} = 3$	$\gamma_G = 2.5$		
Homeostatic Parameters (6.4)			
$h_{max} = 1$	$k_2 = -9$	$\tau_{hw} = 600$ seconds	$h_{th} = 2$
$h_{min} = 0$	$k_1 = 0.5$	$\tau_{hs} = 600$ seconds	
Noise Stimuli Parameters (6.6)			
Width = 1.6	Amplitude = 20		
Noisy h_{max} Parameters (6.7)			
$r_h = 0.003$	Mean = 1	Amplitude = 10	

Table 6.1: **Baseline Parameter Values for two-state flip-flop ODE model with exogenous noise.** α_P and β_P are in units of effective synaptic input. $g_{n,P}$ has units of effective synaptic input/Hz. h_{max} and h_{min} have units of percentage mean slow wave activity (SWA). k_1 and k_2 have units of effective synaptic input and effective synaptic input/(% mean SWA), respectively.

generate this noise and incorporate it to the ODE model.

6.3.3.1 Noise Stimuli to Wake-promoting Population

The wake-promoting regions also receive inputs from higher brain regions including the cortex and the thalamus. These inputs represent peripheral signals such as auditory sensations and motor responses [157]. We model this variability using a noisy stimulus that acts as an input to the wake-promoting neuronal population and modulates its firing rate.

Numerical evidence shows that the wake activation of the light phase experimental data roughly follows a Poisson distribution, see Section 6.2 for details. Consequently, we model the exogenous noise using a Poisson distribution. This results in a spike-like behavior in the firing rate time-series of f_W as shown in Figure 6.6.

To understand this, consider a Poisson event that occurs at a rate of r. Then, we follow Algorithm 6.1 to generate a noise trace of the same size as the time span vector, $\vec{t} = [t_0, t_1, \ldots, t_{n-1}]$, before solving the system of ODEs at these time points with the associated

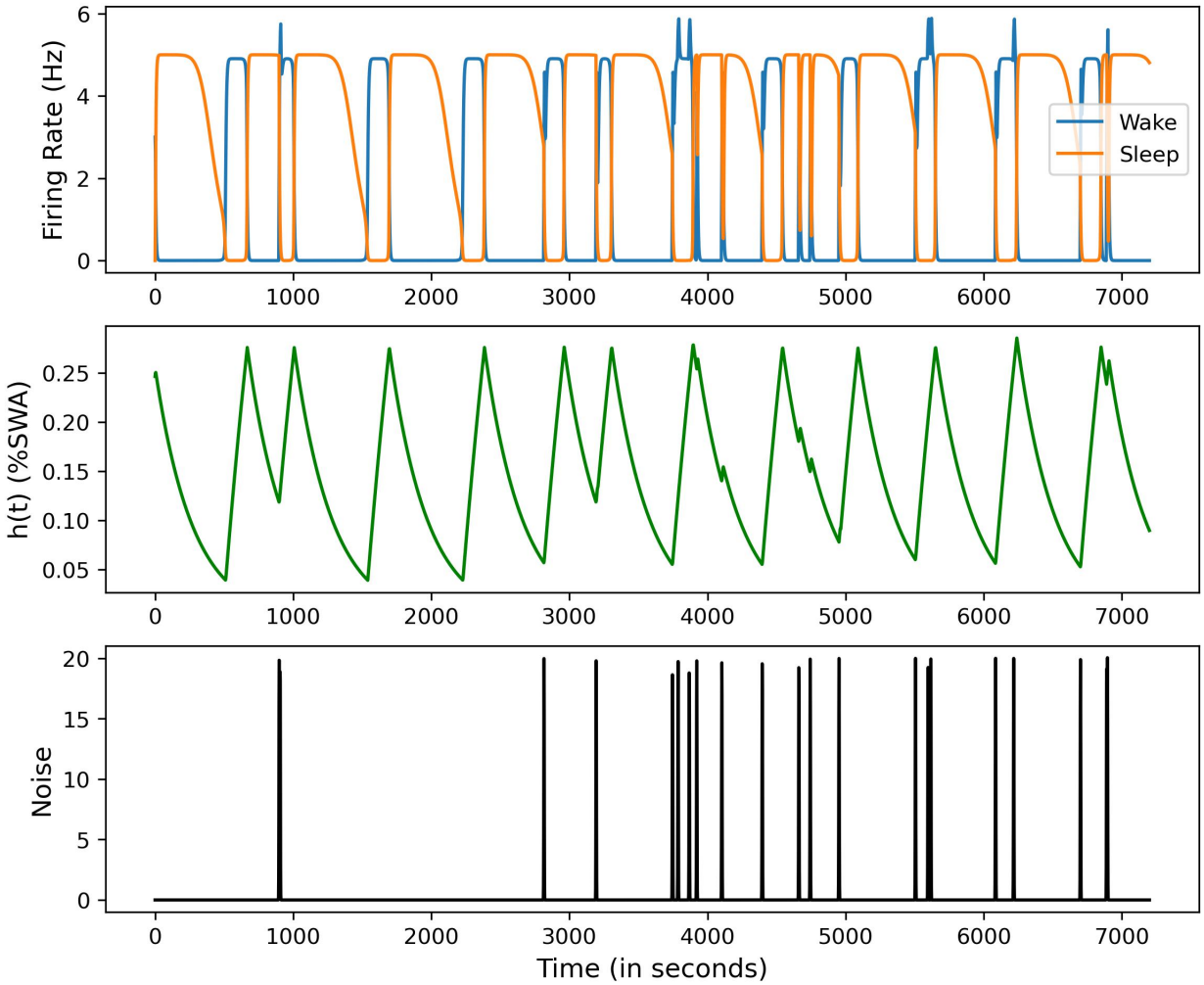


Figure 6.6: **Time traces showing the variation of f_W, f_S, $h(t)$ and Noisy stimuli over 2 hours**. The traces are obtained by numerically solving the ODE system with modified (6.6) and initial conditions $f_W(0) = 3$, $f_S(0) = 0$, $h(0) = 0.24637$. The system is solved for every 1.25 seconds. The value of r is fixed at 0.003.

noisy input.

Finally, we modify equation (6.1a) to include the corresponding noise input,

$$\frac{\mathrm{d}f_W}{\mathrm{d}t}(t_i) = \frac{1}{\tau_W}\left(W_\infty(-g_{G,W}C_G + \text{Noise[i]}) - f_W\right), \tag{6.6}$$

where i denotes the index at which t_i occurs in the time span vector. As shown in Figure 6.6, the new system of ODEs yields noisy and irregular dynamics similar to the experimental

Algorithm 6.1 Generate Noise Stimuli

Input $\vec{t} = [t_0, t_1, \ldots, t_{n-1}]$ of length n
Initialize Noise $= [0, 0, \ldots, 0]$ of length n
Initialize $\tau = 0$; τ denotes the time points for noise stimuli
Initialize Width $= 1.6$; Width denotes the width of a noise stimulus
Initialize Amplitude $= 20$; Amplitude denotes the maximum amplitude of a noise stimulus
while $\tau < t_{n-1}$ **do**
    Generate a random number u uniformly distributed in (0,1)
    $\tau \leftarrow \tau - \frac{1}{r}\log(u)$
    Noise $\leftarrow$ Noise + Amplitude $\times \exp\left(-\frac{(\vec{t}-\tau)^2}{2\times \text{Width}^2}\right)$
end while

data. Importantly, Algorithm 6.1 ensures that the rate parameter r is directly involved in the computation of model output. This will prove essential for the implementation of automatic differentiation of output error with respect to r.

6.3.3.2 Noise for Homeostatic Sleep Drive

Another potential source of stochasticity in the model is the fluctuating h_{max}, a parameter that influences the switch between wake and sleep states [9]. Biologically, it represents the variability in the homeostatic sleep drive. Recall the differential equation governing the accumulation and dissipation of homeostatic sleep drive (6.4). Instead of a fixed parameter h_{max}, we consider a noisy time trace of the same size as $\vec{t}$ generated by Algorithm 6.2.

We modify (6.4) to include a noisy h_{max} input as

$$\frac{\mathrm{d}h}{\mathrm{d}t}(t_i) = \frac{\mathcal{H}(f_W - h_{th}).(h_{max}[i] - h)}{\tau_{hw}} + \frac{\mathcal{H}(h_{th} - f_W).(h_{min} - h)}{\tau_{hs}} \tag{6.7}$$

where i denotes the index at which t_i occurs in the time span vector, $\vec{t}$.

Figure 6.7 illustrates the effect of using a noisy h_{max} trace in the deterministic ODE model. A higher value of h_{max} allows for greater accumulation of homeostatic sleep drive during wakefulness, which can lead to longer sleep bouts. In contrast, a lower h_{max} reduces the ceiling for this accumulation, resulting in weaker sleep pressure, longer wake bouts, and

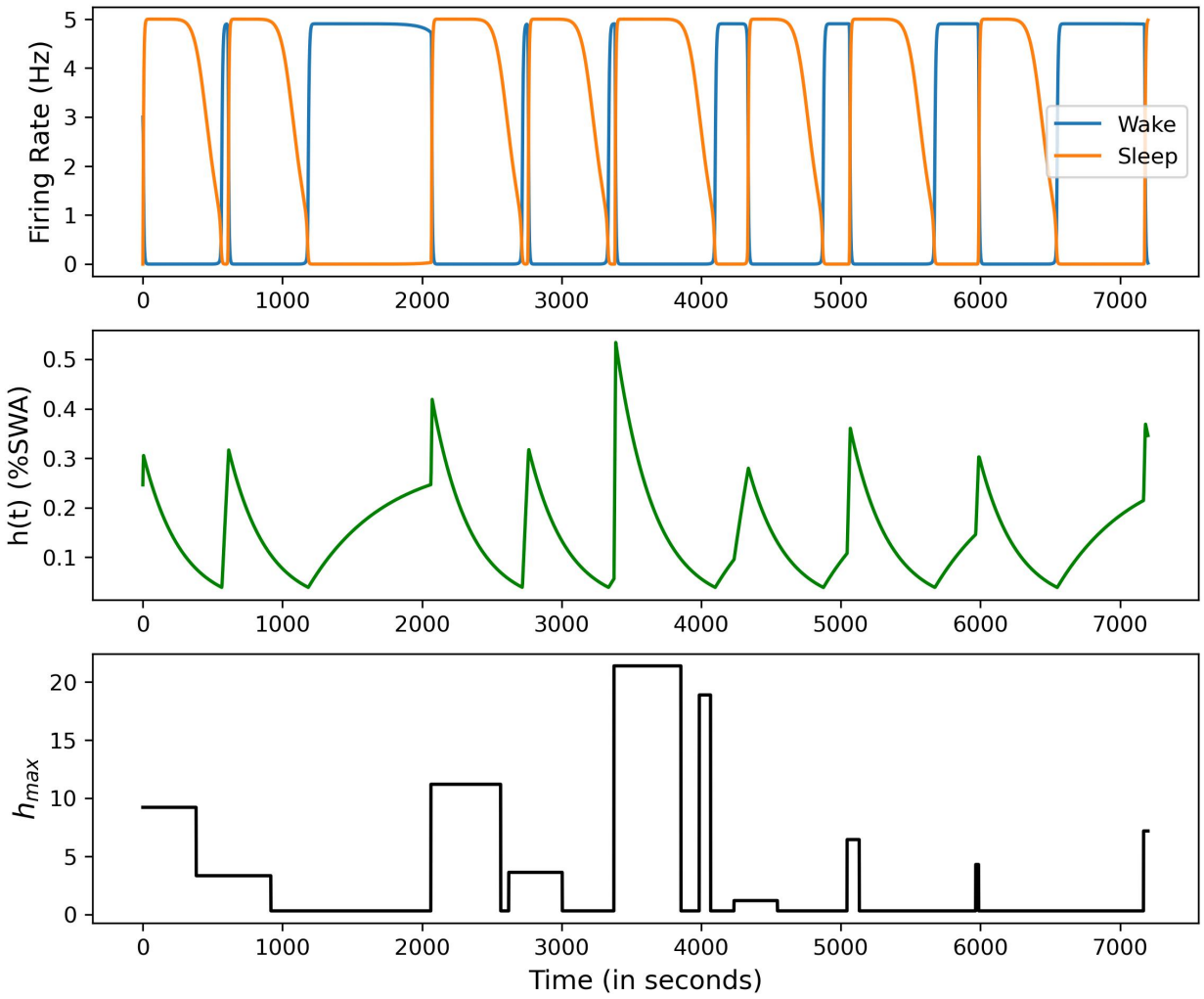


Figure 6.7: **Time traces showing the variation of f_W, f_S, $h(t)$ and Noisy h_{max} over 2 hours.** The traces are obtained by numerically solving the ODE system with modified (6.7) and initial conditions $f_W(0) = 3$, $f_S(0) = 0$, $h(0) = 0.24637$. The system is solved for every 1.25 seconds. The value of r_h is fixed at 0.003, amplitude for $h_{max} = 20$ and mean for $h_{max} = 1$.

shorter sleep episodes compared to the deterministic baseline.

In Algorithm 6.2, we clamp h_{max} at a lower bound of 0.3. This is motivated by the fact that in the non-noisy model, $h(t) = 0.3$ roughly marks the transition threshold between wake and sleep states. Allowing h_{max} to fall below this value could result in nonphysiological dynamics where the model remains trapped in the wake state.

Importantly, this noise component contributes subtle variability to both wake and sleep

Algorithm 6.2 Generate Noisy h_{max}

Input $\vec{t} = [t_0, t_1, \ldots, t_{n-1}]$ of length n
Initialize $\vec{h}_{max} = [0, 0, \ldots, 0]$ of length n
Initialize Amplitude_$h_{max} = 10$; Mean_$h_{max} = 1$
Generate a random number u uniformly distributed in (0,1)
Initialize $t_{start} = 0$ and $t_{end} = -\frac{1}{r_h}\log(u)$
while $t_{start} \leq t_{n-1}$ **do**
    Generate a random number z normally distributed with mean 0 and variance 1
    $h_{max} \leftarrow \max((\text{Mean}_h_{max} + z \times \text{Amplitude}_h_{max}), 0.3)$
    Let $I = \{i \mid t_{\text{start}} \leq \vec{t}[i] \leq \min(t_{\text{end}}, t_{n-1})\}$
    $\vec{h}_{max}[I] = h_{max}$
    Generate a random number u uniformly distributed in (0,1)
    $t_{start} \leftarrow t_{end}$
    $t_{end} \leftarrow t_{end} - \frac{1}{r_h}\log(u)$
end while

durations and yields bouts that are slightly shorter or longer than those produced by the deterministic model. This is in contrast to the noise added to the wake-promoting population, which tends to fragment the sleep-wake dynamics by inducing very short wake bouts. As we will demonstrate in Section 6.5, both types of noise are essential to reproduce the variability and distributional features observed in the experimental data.

6.4 Methodology: Fitting Model to Data

In the previous sections, we visualized the experimental data and analyzed the structure and behavior of the noisy mathematical model. In this section, we outline the methodology for efficiently fitting the model to the data using a gradient-based optimization framework.

We begin by detailing the components of the implemented gradient descent algorithm. Next, to better understand the influence of various model parameters, we present a simplistic sensitivity analysis using the deterministic model. We also introduce a strategy for augmenting the experimental data and processing model outputs to enhance optimization performance. Finally, we explicate the algorithm decisions of the optimization framework, including the definition of the error function and the evaluation metrics.

6.4.1 Gradient Descent using Automatic Differentiation

Optimizing a noisy model to fit experimental data presents significant challenges. In particular, the presence of exogenous noise precludes the use of straightforward error metrics, such as absolute differences or root mean squared error (RMSE), to directly compare the model output to the experimental data at specific time points.

Previous works have employed a variety of optimization techniques for similar tasks, spanning Bayesian inference, simulation-based inference, and neural network approaches [178, 175, 149, 151]. Bayesian methods, in particular, can be subdivided into two categories: those that assume a known likelihood function for the experimental data, and those that operate without explicit knowledge of the likelihood. Despite the successes of these approaches, a major limitation is their reliance on specifying an informed prior distribution over the model parameters. In practice, obtaining reliable priors can be difficult, especially for complex biological systems where limited prior knowledge is available. Consequently, relatively limited work has addressed parameter estimation without assuming confident knowledge of parameter priors.

To circumvent these issues, we implement a gradient-based optimization algorithm that does not require prior distributions or likelihood functions. However, the structure of our model output makes this approach non-trivial. Since the model generates noisy trajectories, we encounter two major challenges. First, the choice of an appropriate error function that accurately captures the differences between noisy model outputs and noisy temporal data becomes critical. Second, the model's gradients with respect to parameters cannot be derived explicitly and hence an automated approach to gradient computation is warranted.

For this task, we conceptualize each experimental data instance (such as that illustrated in Figure 6.2) as a sample drawn from an underlying distribution. For the two-state model, the trajectory denoting the firing rate of the wake promoting population basically captures both the wake and sleep dynamics over time. We consider the f_W time trace as a sample of

our model output.[3]

Accordingly, the objective becomes to estimate the model parameters such that the model outputs samples from the same distribution as the experimental data. This formulation aligns with the modeling philosophy for sleep-wake regulation detailed in Section 6.3, where the system exhibits deterministic dynamics perturbed by external noise. Thus, optimization involves minimizing a distributional discrepancy between model-generated samples and data samples, under the assumption that both originate from a shared underlying structure.

The key steps of the gradient-based optimization algorithm are as follows. For a given set of parameter values, we generate N independent samples from the noisy ODE model. We then compute an error metric that quantifies the discrepancy between the N model samples and N experimental data samples. The number of samples is kept identical to avoid statistical skewness. Finally, we compute the gradient of the error metric with respect to the model parameters, and update the parameters accordingly for the next iteration. These steps are summarized schematically in Algorithm 6.3.

Algorithm 6.3 Gradient-Based Optimization Algorithm

Input fixed parameters, initial conditions and step size to numerically solve ODE model
Input N target samples from experimental data
Input initial parameter values θ, learning rates corresponding to θ and number of optimization iterations $\mathcal{N}$
while $\mathcal{N} > 0$ **do**
    Collect N samples from model output at θ
    Error = Criterion(Current Samples, Target Samples)
    Compute gradient of the error with respect to θ
    Update θ using an optimizer
    $\mathcal{N} \leftarrow \mathcal{N} - 1$
end while

Since the model output is stochastic and implicitly defined through numerical ODE solutions, there is no straightforward way to analytically compute the required gradients. To address this, we utilize the framework of automatic differentiation, a technique widely adopted

[3]We apply a normalization step on the f_W time trace to match it to the range of sleep scores recorded in the experimental data i.e., between 0 to 1. See Section 6.4.3 for more details.

in deep learning and scientific computing pipelines. Automatic differentiation operates by essentially tracking the sequence of arithmetic operations during the forward evaluation of the model, thus forming a computational graph. Importantly, the optimization parameters must be explicitly included in the computational graph; otherwise, the gradients with respect to these parameters cannot be computed. Additionally, care must be taken to avoid breaking the computational graph through discrete operations, such as if-else conditions, which can disrupt the flow of derivative information. It then applies the chain rule systematically to back-propagate through these operations, yielding exact gradients at the current parameter values without symbolic derivation or finite differences.

Additional details regarding pre-processing steps, the choice of error metric and optimization function, and specific algorithmic design choices are provided in subsequent sections. These components are instrumental in building an efficient optimization pipeline, from understanding parameter sensitivities to evaluating model performance.

6.4.2 Sensitivity Analysis for Model Parameters

One of the most critical steps in an optimization pipeline is selecting a robust set of parameters to estimate. Ideally, the parameters should significantly influence the model output and exhibit low mutual correlation to avoid non-identifiability and the risk of multiple local minima in the error landscape. As a first step in our analysis, we perform a preliminary sensitivity analysis using the baseline deterministic model to evaluate the involved parameters.

In the deterministic flip-flop model, the primary outputs of interest are wake durations and sleep durations (or inter-wake intervals). To assess sensitivity, we perturb each baseline deterministic parameter by $\pm 20\%$ and observe the resulting changes in these durations using pairwise parameter plots. The perturbation criterion is chosen to balance capturing nonlinear sensitivities while remaining within a biologically reasonable parameter range. For reference, Figures 6.8 and 6.9 show the pairwise plots of k_2 and k_1 with other parameters. Similar pairwise relationships are observed in the remaining set of pairwise plots.

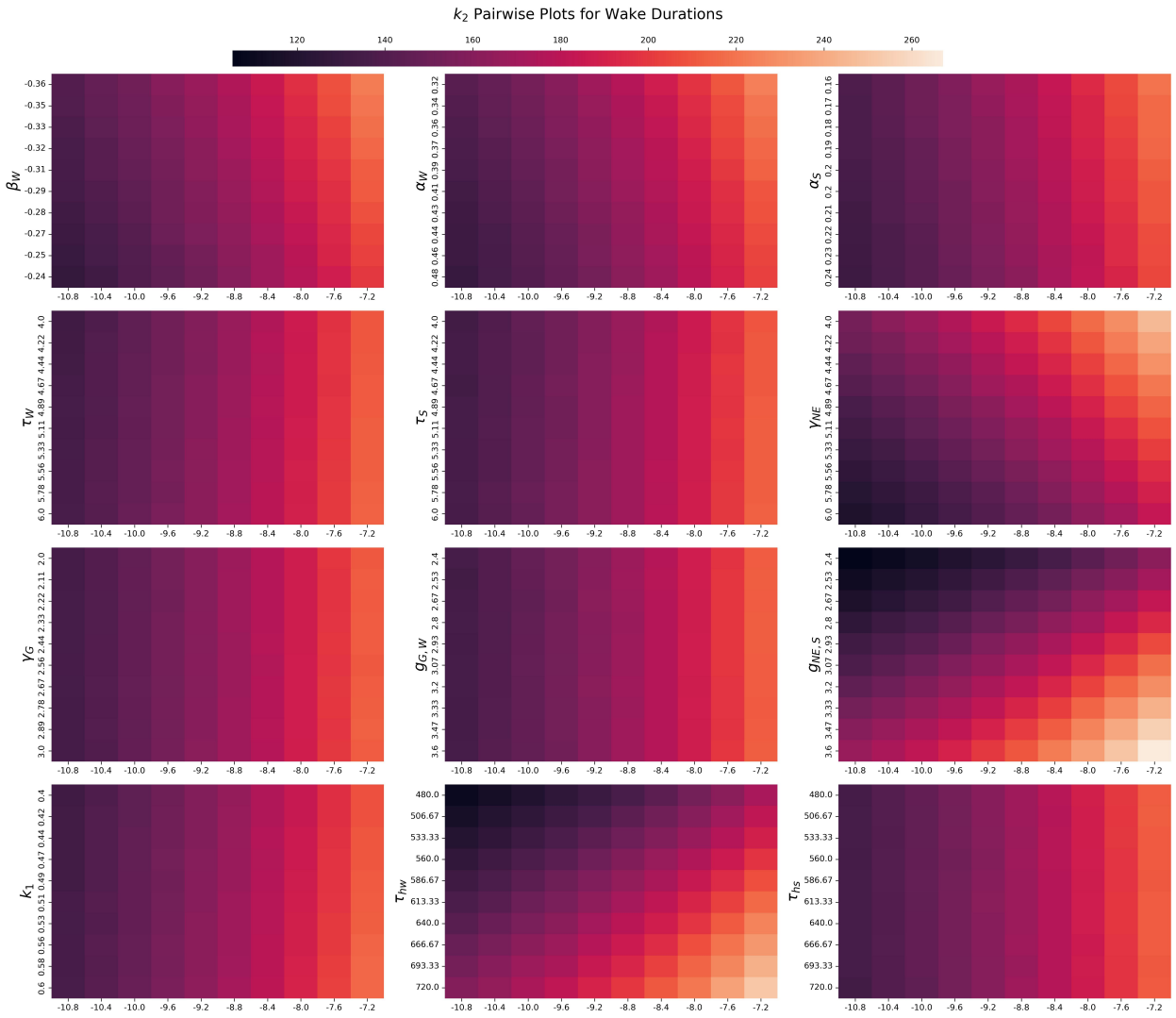


Figure 6.8: **Pairwise plots between k_2 and other model parameters show correlated wake duration variation in the deterministic model.** We consider a $\pm 20\%$ perturbation from the baseline values of the parameters.

We find that γ_{NE}, $g_{NE,S}$, k_2 and τ_{hw} are highly correlated to each other and correspond to a significant change in the wake durations of the model output. Similarly, γ_{NE}, $g_{NE,S}$, k_1 and τ_{hs} exhibit correlated and sensitive effects on inter-wake intervals. These correlations are evident through consistent diagonal or anti-diagonal gradients in the pairwise plots. The parameters β_W, α_W and α_S show weak correlation to the set of strongly correlated parameters stated above. By contrast, the parameters such as τ_W, τ_S, γ_G and $g_{G,W}$ have minimal impact on the wake durations and inter-wake durations. In pairwise plots, these

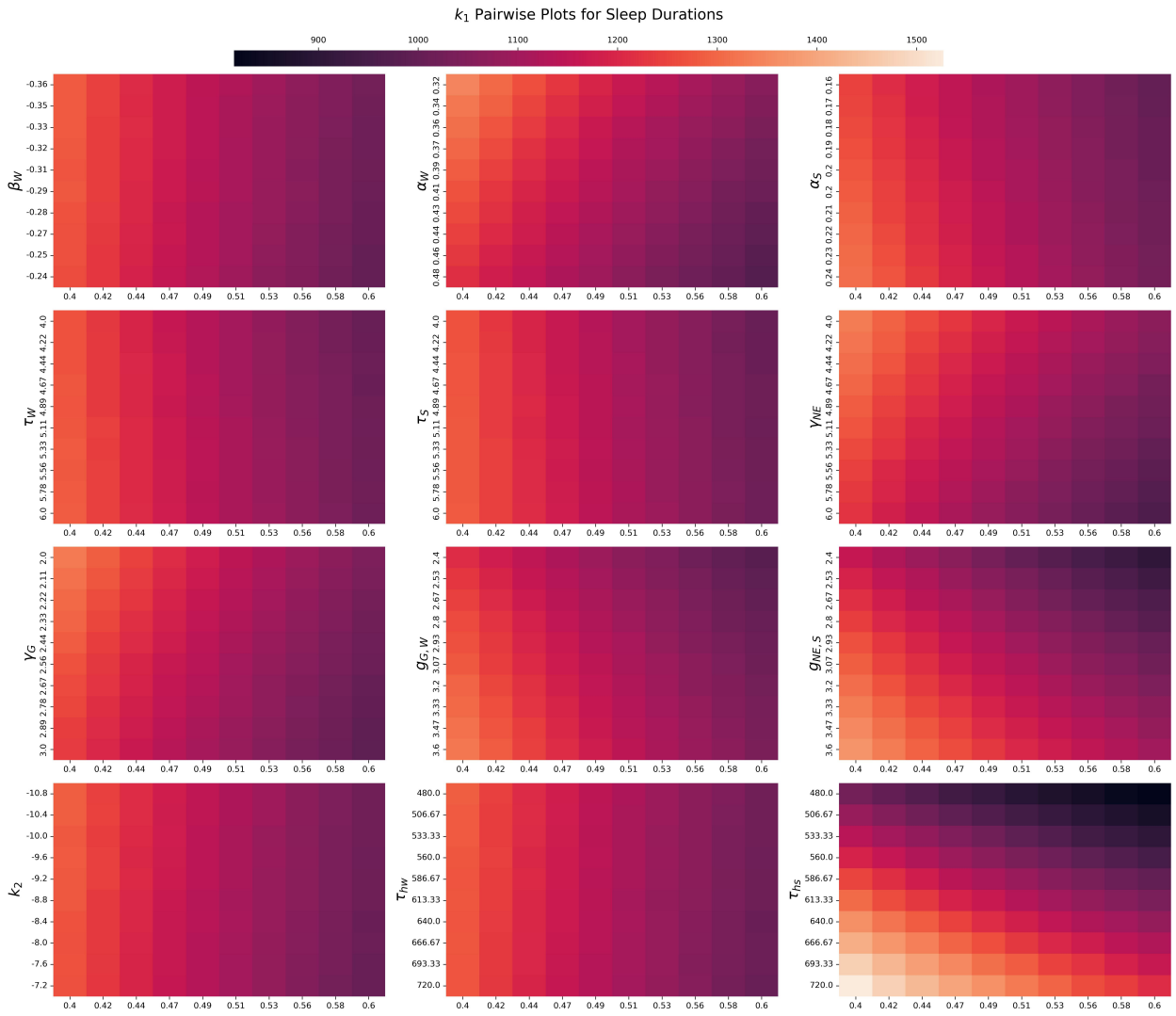


Figure 6.9: **Pairwise plots between k_1 and other model parameters show correlated sleep duration variation in the deterministic model.** We consider a $\pm 20\%$ perturbation from the baseline values of the parameters.

less sensitive parameters are effectively *dominated* by the strongly influential and correlated parameters.

Several conclusions follow from this analysis. First, including multiple highly correlated parameters in the optimization could impair performance by introducing multiple local minima. Conversely, including parameters with minimal sensitivity could be inefficient, as their gradients would be trivial compared to those of dominant parameters and may hinder optimization progress.

Thus, selecting the final set of optimization parameters requires careful reasoning, especially given that the introduction of noise alters the system's dynamics. Pairwise relationships observed in the deterministic model may not carry over directly to the noisy model. Furthermore, pairwise compensation in the error landscape does not necessarily imply compensation across the full parameter space; a global minimum for the full set could still exist despite local minima evident in lower-dimensional slices. We discuss the final selection of optimization parameters in more detail in Section 6.5.

6.4.3 Data Processing and Augmentation

The performance of a gradient-based optimization algorithm critically depends on the structure and quality of the data being used. To ensure that the model outputs resemble the experimental data without altering the underlying biological relevance, we apply a smooth normalization step. Specifically, the normalized time series traces are mapped to the interval [0, 1], featuring sharp transitions to resemble the discrete sleep scores observed experimentally. The transformation is given by

$$\text{Normalized Sample} = \frac{2}{1 + \exp(-\chi * \text{Sample})} - 1 \tag{6.8}$$

where χ is fixed at a value of 0.5. This normalization maintains the differentiability of the output while avoiding explicit discretization, a critical requirement for preserving the computational graph necessary for automatic differentiation. Importantly, no modifications are applied to the experimental data itself.

As mentioned earlier, it is good practice to sample an equal number of data points from both the model output and the experimental data to avoid biasing the discrepancy measure. We select 100 samples per optimization iteration to ensure stable estimates of the distributional discrepancy. However, the available experimental recordings in the light phase only provide 50 instances (see Section 6.2), necessitating a simple data augmentation strategy.

To augment the light phase dataset, we partition the available observations into two non-

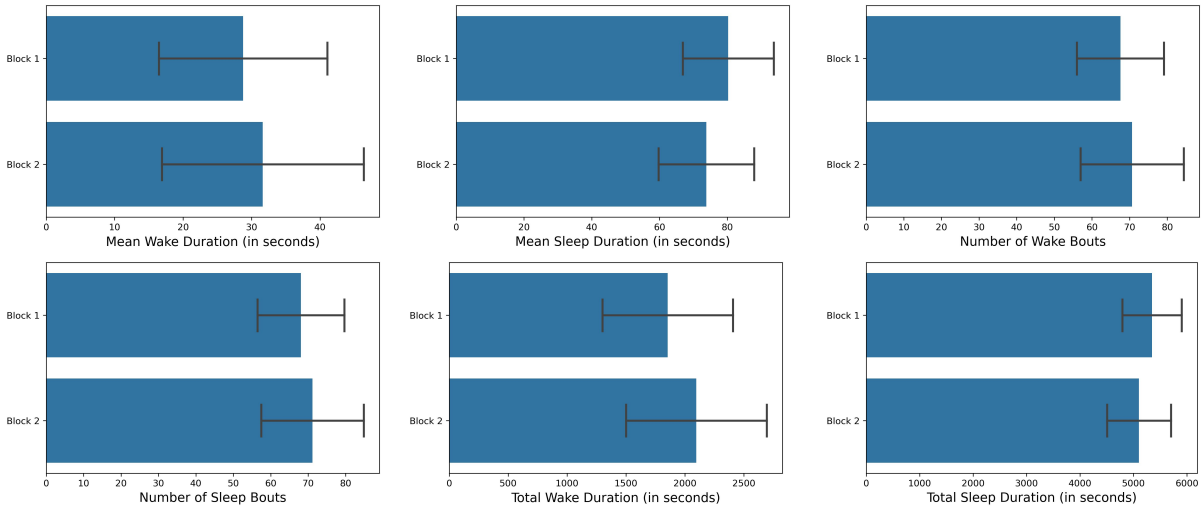


Figure 6.10: **The six inference measures show similarity over the two hour time blocks.** Block 1 corresponds to 50 mice recordings from 9 AM to 11 AM. Block 2 corresponds to 50 mice recordings from 11 AM to 1 PM. See C.1 for sample-level details.

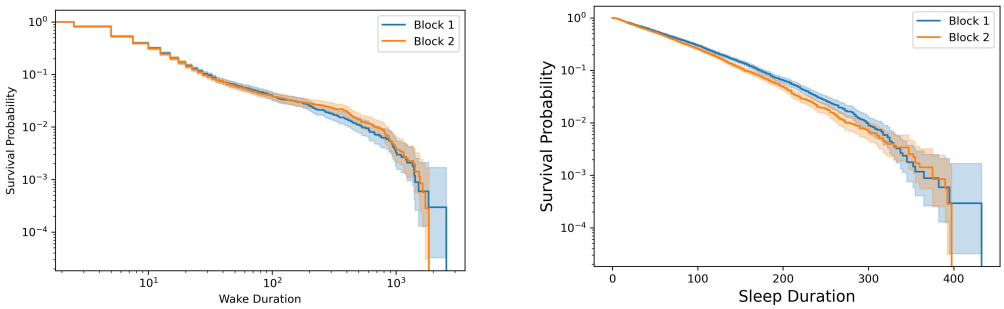


Figure 6.11: **Survival curves overlap for wake and sleep durations collected for light phase data over 9 AM-11AM and 11AM-1PM.** This motivates the augmentation of the two sets to collate a light phase dataset of 100 mice recordings.

overlapping two-hour blocks: 9 AM to 11 AM and 11 AM to 1 PM. The samples from the second block are treated as statistically independent, allowing us to append them to the first block and reach the target of 100 instances. To validate this approach, we plot the six descriptive measures for both time blocks and observe close agreement using the Mann-Whitney U test (p-value > 0.01 for all measures) [116], as shown in Figure 6.10. Figure 6.11 compares survival curves of wake and sleep durations collected over 100 samples of data and model simulations with different optimization choices. The significant similarity in dynamics between the two intervals supports the assumption that both sample sets arise from the same

underlying distribution.

Remark 6.1. In this chapter, we use survival curves to compare the distributions of sleep and wake bouts by aggregating all durations observed over the 100 samples. We plot the survival curves for wake durations on a log-log scale and the survival curves for sleep durations on a semi-log scale, consistent with other literature [13, 9].

6.4.4 Choice of Optimization Hyperparameters: Error, Kernel and Optimizer

Selecting an appropriate discrepancy measure between the model output and the experimental data is another critical step for the success of the optimization pipeline. Among the commonly considered options are the Maximum Mean Discrepancy (MMD), Wasserstein distance, and Kullback-Leibler Divergence (KLD). Here, we justify the choice of MMD based on several practical and theoretical considerations relevant to our setting.

Firstly, the Kullback-Leibler divergence, while attractive in theoretical contexts - particularly due to its close relationship with log-likelihood and its favorable properties for models from the exponential family, requires explicit knowledge of the underlying probability density functions (PDFs) for both the data and the model output [107]. In our case, we only have access to sampled data without an explicit analytical form for the PDFs, rendering KL divergence infeasible.

The Wasserstein distance, on the other hand, can be computed directly from samples and is known for its ability to capture differences between distributions [143]. Theoretically, it is computed as

$$\mathrm{WD}(p,q) = \inf_{\pi \in \Pi(p,q)} \int_{\mathbb{R}\times\mathbb{R}} |x-y| \mathrm{d}\pi(x,y) \tag{6.9}$$

where $\Pi(p,q)$ is a set of probability distributions on 1-D probability mass functions, p and q. The empirical computation of Wasserstein distance can be expensive, especially when numerous samples of the data are involved. For our setting, computing it over 100 time trace samples for each optimization iteration might lead to extensive overhead. Instead, we utilize this measure to evaluate the model performance on the population-level by collating

wake and sleep bout durations as 1-D sample sets from the model outputs and data.

In contrast, MMD, first introduced by Gretton et al. [70], provides an efficient and practical alternative. It operates by evaluating difference in means of kernel embeddings of the respective samples, thus avoiding costly optimization procedures. Empirically, MMD translates to

$$\text{MMD(P,Q)} = \left(\frac{1}{m_1^2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_1} k(x_i, x_j) - \frac{2}{m_1 m_2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} k(x_i, y_j) + \frac{1}{m_2^2} \sum_{i=1}^{m_2} \sum_{j=1}^{m_2} k(y_i, y_j) \right)^{1/2} \tag{6.10}$$

where P and Q refer to distributions and $\{x_1, \ldots, x_{m_1}\}$ and $\{y_1, \ldots, y_{m_2}\}$ refer to their respective sets of samples [70].

MMD depends on the choice of kernel function, and identifying an appropriate kernel is often non-trivial. Gaussian kernels are one of the most popular kernel metrics and offers benefits such as smoothness and computational efficiency, essential for our optimization pipeline [111, 189]. For simplicity, we implement the MMD measure as our error criterion using the Gaussian kernel for our current analysis. For two samples x and y, the Gaussian kernel is defined as

$$k(x, y) = \exp\left(-\frac{||x - y||^2}{n} \right) \tag{6.11}$$

where n denotes the length of x and y.

We utilize the Adaptive Moment Estimation (ADAM) optimizer to update the parameter values [102]. The ADAM optimizer performs better than the vanilla gradient descent algorithm as it takes advantage of two key modifications: Momentum and Root Mean Square Propagation (RMSprop). These help in accelerating the gradient descent by dynamically adapting the learning rate and memorizing the past gradients for better convergence.

6.4.5 Model Evaluation

Finally, we describe the procedure used to evaluate the performance of the optimization algorithm. In the context of high-dimensional parameter spaces and noisy data, the algorithm may not converge to a single set of parameter values. Instead, multiple distinct parameter sets may yield similarly low error values and generate dynamics comparable to the experimental data. Therefore, it is critical to establish a robust inference pipeline to assess convergence and model performance.

For each optimization run, we compute the average and standard deviation (std) of the converged parameter values over the final set of iterations. The number of iterations included in this analysis depends on the total length of the optimization run. For example, if an experiment runs for 300 iterations, we typically use the last 50 iterations to estimate the mean and variability of the inferred parameters.

To assess the quality of the model fit, we evaluate the error metric, the MMD value, at the optimized parameters. A natural question arises: how do we interpret a given MMD value, and when is it "small enough"? To contextualize the MMD value, we computed the MMD between two disjoint subsets of the experimental data: 50 random samples versus the remaining 50. This yielded an average value of approximately 0.0775 over 5 runs. This MMD value serves as a rough lower bound on the discrepancy between distributions arising purely from within-sample variability for 50 samples. As the sample size increases, this empirical MMD decreases and approaches the true divergence between the underlying distributions.

Additionally, we compare the six key inference measures computed from the model output at the averaged parameter values to measures computed from the experimental data.. We also include error bars, defined by average $\pm$ std, to capture the uncertainty around these estimates for a sample-level evaluation. In addition, we generate survival curves, histogram plots and compute the Wasserstein distance for sleep and wake durations to provide a comprehensive comparison between the model-generated and experimental population-level distributions. The results and their implications are discussed in detail in the upcoming

Maximum Mean Discrepancy	**Optimized** r	**+ Optimized** (τ_{hw}, τ_{hs})	**+ Optimized Weak Parameters**
With Noisy Stimuli	0.1665	0.1124	0.1115
With Noisy Stimuli + Noisy h_{max}	0.1469	0.1106	0.1031

Table 6.2: **Maximum Mean Discrepancy improves as we utilize noisy h_{max} for the optimized models.** The values denote the MMD measure computed between 100 data samples and 100 model outputs. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameters. Optimized Weak Parameters refers to model with (6.12), (6.13) and (6.14) and other baseline parameters.

Section 6.5.

6.5 Results

To begin our analysis, we first identify parameter sets that both capture the observed sleep-wake patterns and ensure parameter identifiability for gradient-based optimization. Our approach to incorporating noise into the deterministic flip-flop model proceeds in two stages: first, by adding exogenous noise to the wake population's firing rate; second, by introducing a noisy time trace for the homeostatic input ceiling, h_{max}.

With noise added, the rate for noise stimuli r becomes a key driver of stochastic sleep-wake behavior, making it a primary focus of our optimization. A second set of influential parameters are the time constants τ_{hw} and τ_{hs} which govern the rise and decay of the homeostatic sleep drive $h(t)$ respectively. These constants play a critical role in determining sleep-wake transitions in the deterministic model and are well-studied in the sleep literature for humans [150, 3] and rodents [147, 81, 59, 182]. Further, sensitivity analysis on the deterministic model shows the strong influence of τ_{hw} on wake durations and of τ_{hs} on sleep durations, refer to Section 6.4.2.

To avoid parameter compensation, we avoid including other highly correlated parameters such as γ_{NE}, $g_{NE,S}$, k_2 and k_1 in our analysis. Instead, we focus on a subset of weakly cor-

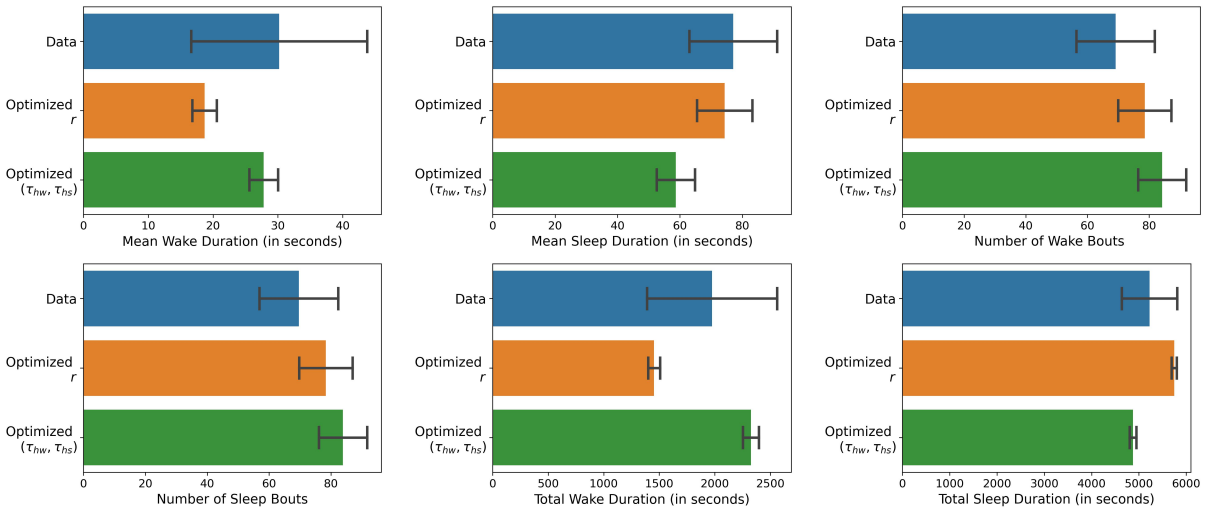


Figure 6.12: **Inference Measures show some discrepancies for optimized model with noisy stimuli input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameters.

related but physiologically relevant parameters: β_W, α_W, α_S and $g_{G,W}$.[4] These parameters affect the firing rate dynamics for both wake and sleep populations and mathematically represent their excitability and synaptic connectivity. In sleep research, such markers are critical for characterizing changes in brain function associated with aging and disease progression [110].

We implement the two strategies for noise addition and evaluate their optimization performance for the parameter sets in the forthcoming subsections.

6.5.1 Model Dynamics are Driven by f_W Noise Stimuli and the Underlying Deterministic Behavior

As the first step, we incorporate a noisy stimuli input to the firing rate of the wake-promoting population as described in Section 6.3.3. In this experiment, the rate of the noise stimulus is optimized, while all other model parameters are held fixed at their baseline values (see Table 6.1). After 300 optimization iterations, the noise rate over the last 50 iterations

[4]Parameters τ_W and τ_S have minimal impact on wake and sleep durations and mostly influence the amplitude of f_W and f_S. Parameter γ_G is often strongly correlated to $g_{G,W}$.

converges to:

$$r = 0.01699 \pm 0.00021. \tag{6.12}$$

The Maximum Mean Discrepancy (MMD) for the model output at this average parameter value is 0.1665.[5] Notably, the optimized value of r is close to the Poisson rate estimated from the empirical distribution of sleep durations ($r = 0.0127$), supporting the choice of noise generating algorithm (see Section 6.2).

To further refine the model, we optimize the time constants τ_{hw} and τ_{hs}. We do not jointly optimize all three parameters because the rate r strongly influences system dynamics and tends to dominate other parameters in an optimization pipeline. Typically, the gradient of the error with respect to other parameters is trivial compared to the gradient with respect to r. Hence, we fix r at its converged value (6.12) and perform a second round of optimization.

The time constants converged as follows:

$$\tau_{hw} = 553.26234 \pm 2.08506 \text{ seconds}, \tag{6.13a}$$

$$\tau_{hs} = 273.28565 \pm 1.85658 \text{ seconds}, \tag{6.13b}$$

with a corresponding MMD of 0.1124 at the average parameter values. These time constants are lower than their baseline counterparts and result in an improved distributional match between model and data as indicated by the lower MMD.

To evaluate the model performance, we revisit the six inference measures introduced in Section 6.2. These measures - mean and total durations, and bout counts for wake and sleep - capture the *sample-level variability* of the model output across individual traces. We refer to the results of the first experiment with varying r as 'Optimized r' and the results of the second experiment with varying (τ_{hw}, τ_{hs}) on the converged value of r as 'Optimized (τ_{hw}, τ_{hs})'.

[5]For completeness, the MMD at the lower and upper ends of the converged range of r are 0.1841 and 0.1517, respectively. We do not consider the converged value of MMD from the optimization runs as it might not be a reliable measure of the MMD at the converged value of parameters.

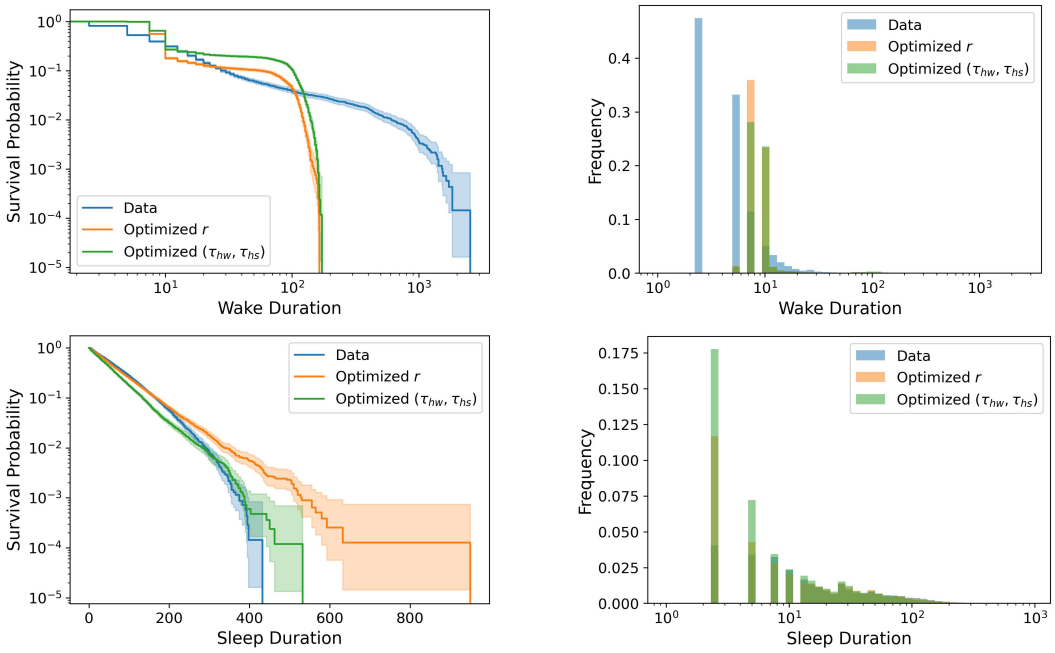


Figure 6.13: **Population-level survival curves and histograms show some discrepancies for optimized model with noisy stimuli input.** While the shape of the sleep distribution is captured by the model output, the wake distribution lacks the elongated shape admitted by the data. Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameters.

Figure 6.12 shows strong alignment in the means and standard deviations of these metrics between the model and data.[6] However, the optimized model slightly overestimates the number of wake and sleep bouts, with corresponding decreases in their average durations. This points to an increased fragmentation of wake and sleep states, likely induced by the noise perturbing the underlying deterministic dynamics.

To further understand this behavior, we examine *population-level distributions* of wake and sleep bouts. By pooling bout durations across all 100 simulated traces, we construct survival curves and log-binned histograms, as shown in Figure 6.13. These allow us to compare the full empirical distributions of sleep-wake architecture between the data and model outputs.

[6]While the central tendencies and variances visually overlap, formal statistical comparisons (e.g., Mann-Whitney U test) suggest that some differences remain, particularly in the structure and frequency of bouts.

Wasserstein Distance - Wake	**Optimized r**	**+ Optimized (τ_{hw}, τ_{hs})**	**+ Optimized Weak Parameters**
With Noisy Stimuli	18.7857	25.2335	24.9998
With Noisy Stimuli + Noisy h_{max}	14.1797	11.3886	10.2125
Wasserstein Distance - Sleep	**Optimized r**	**+ Optimized (τ_{hw}, τ_{hs})**	**+ Optimized Weak Parameters**
With Noisy Stimuli	6.4290	17.1438	18.1298
With Noisy Stimuli + Noisy h_{max}	14.9095	16.1211	15.5800

Table 6.3: **Wasserstein distance improves as we utilize noisy h_{max} for our converged models.** The values denote the Wasserstein distance computed between the collective wake and sleep durations observed for 100 data samples versus 100 model outputs. Optimized r refers to the model with r given by (6.12) and other baseline parameters. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13) and other baseline parameter. Optimized Weak Parameters refers to model with (6.12), (6.13) and (6.14) and other baseline parameters.

Interestingly, the wake bout distributions from the model reveal two distinct clusters: one comprising very short wake bouts and another corresponding to much longer bouts. This bimodal pattern reflects a combination of noise-driven transitions (short bouts) and the underlying deterministic attractor (long bouts). In contrast, the empirical data display a more continuous distribution of wake durations, with a significant portion of bouts in the intermediate range (10–500 seconds), which the model fails to capture. Sleep durations in the model are also biased toward shorter lengths, again consistent with high-frequency transitions induced by the noise.

Remark. The survival curves also indicate that the shortest wake bouts generated by the model are longer than those observed in the experimental data. In particular, the rodent sleep-wake dataset includes wake bouts as short as 2.5 seconds. In our current modeling pipeline, we integrate the system of ODEs using the first-order Euler method with a fixed time step of 1.25 seconds. The model output is then downsampled and saved at 2.5-second intervals to match the temporal resolution of the data. This discretization approach limits the ability of the noise input to induce very brief wake bouts lasting only a single bin (2.5 seconds). Addressing this limitation, either through finer time resolution or a more flexible

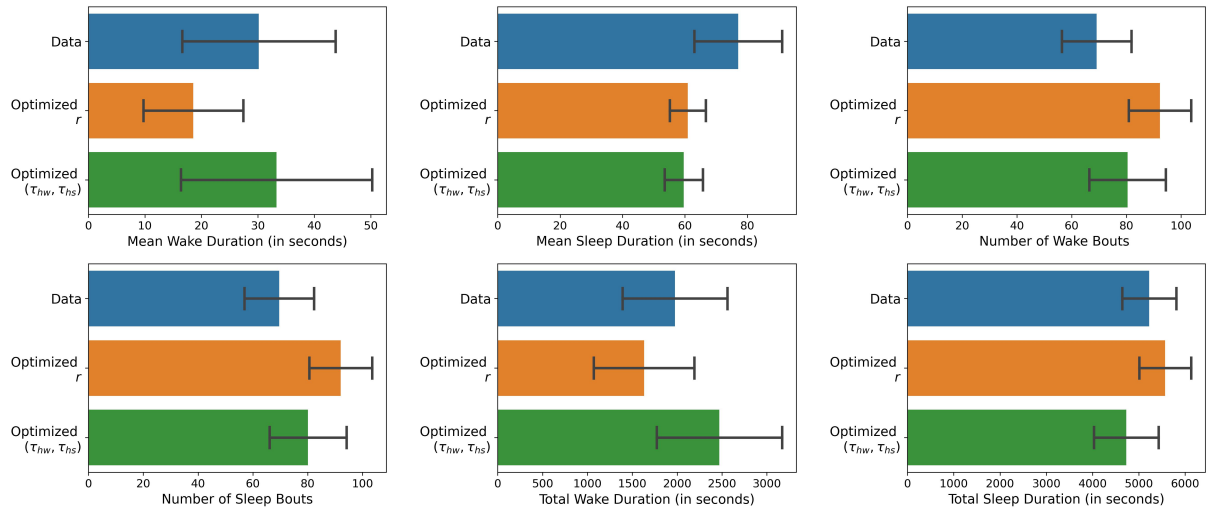


Figure 6.14: **Inference Measures align better and admit more sample-level variability for optimized model with noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12), other baseline parameters with appended noisy h_{max}. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13), other baseline parameters with appended noisy h_{max}.

numerical scheme, will be a key direction for future work.

In summary, the introduction of noise to the wake-promoting population introduces over-fragmentation in the sleep-wake regulation structure. These limitations motivate the need to incorporate stochasticity into other components of the model, such as a noisy h_{max}, as discussed in Section 6.3.3.

6.5.2 Refining Model Fit with Noisy $h_{\max}$ Inputs

To further improve model performance at the population level, we introduce a noisy trace for the $h_{\max}$ parameter, while keeping the optimization-converged values of other parameters fixed. Importantly, incorporating this noise component during the optimization process can obscure the relative contributions of parameters like τ_{hw} and τ_{hs}, due to compensatory dynamics. Therefore, we instead append noisy $h_{\max}$ to models post-optimization, allowing for clearer and more intuitive interpretation of its effect.

When using noisy $h_{\max}$ generated according to Algorithm 6.2, we observe improvements in distributional similarity. For example, in the model with only optimized r, the MMD

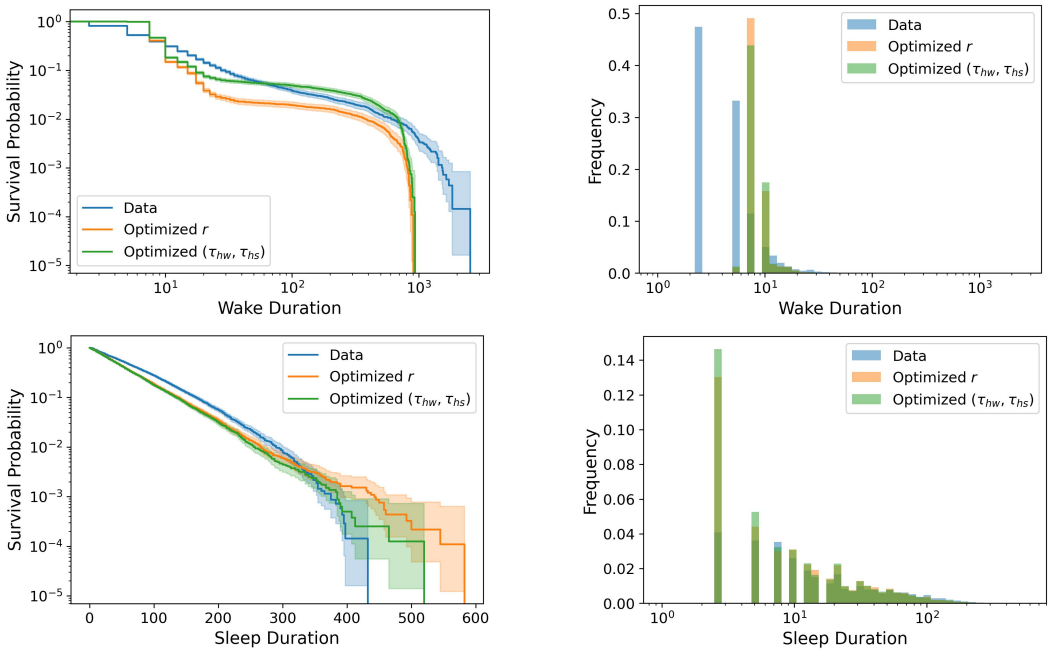


Figure 6.15: **Population-level survival curves and histograms admit similar shapes for optimized model with noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized r refers to the model with r given by (6.12), other baseline parameters with appended noisy h_{max}. Optimized (τ_{hw}, τ_{hs}) refers to model with both (6.12) and (6.13), other baseline parameters with appended noisy h_{max}.

decreases from 0.1665 to 0.1469. For the model with all three parameters optimized (r, τ_{hw}, and τ_{hs}), the MMD decreases from 0.1124 to 0.1106. These values are also summarized in Table 6.2.

The primary impact of noisy $h_{\max}$ is an elongation of the survival curves for wake bout durations, filling in the 10–500 second range that was underrepresented in earlier model variants. This yields a closer match to the empirical wake distribution seen in the data. In parallel, we observe a better visual alignment of sample-level inference measures (see Figure 6.12) between the model outputs and experimental data. For instance, the Mann–Whitney U test for mean wake duration suggests that the model with noisy $h_{\max}$ and optimized $(r, \tau_{hw}, \tau_{hs})$ more closely approximates the empirical mean wake duration values (p-value ~ 0.19).

To further quantify these improvements at the population level, we also compute the Wasserstein distance between model and data distributions. This distance likewise decreases

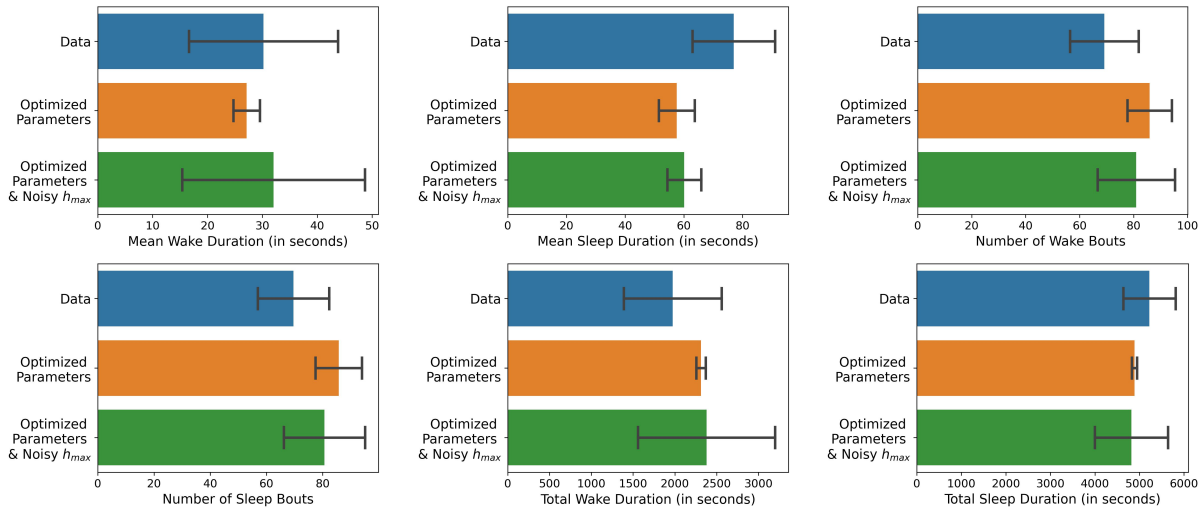


Figure 6.16: **Sample-level similarities between data and model improve for optimized model with both noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized Parameters denote the model with (6.12), (6.13), (6.14) and noisy stimuli. Optimized Parameters & Noisy h_{max} denotes the fully optimized noisy model with appended noisy h_{max}.

following the introduction of noisy $h_{\max}$, supporting the conclusion that the model better captures the overall distributional structure of the sleep-wake architecture for the data (see Table 6.3).

While the updated model captures the general shape of wake and sleep distributions more faithfully, limitations remain in its ability to match the frequency of very short and very long bouts. These extremes are particularly challenging to capture due to inherent biological variability resulting from factors not represented in the model and data noise. Further improvement in the model performance would warrant extensions to the model as well as more finetuning and adjustments both in the baseline model and the optimization framework. The last step we implement is using the converged values of r and the time constants and optimizing a set of less dominant, weakly correlated parameters - β_W, α_W, α_S and $g_{G,W}$ - to assess whether these can further reduce residual discrepancies.

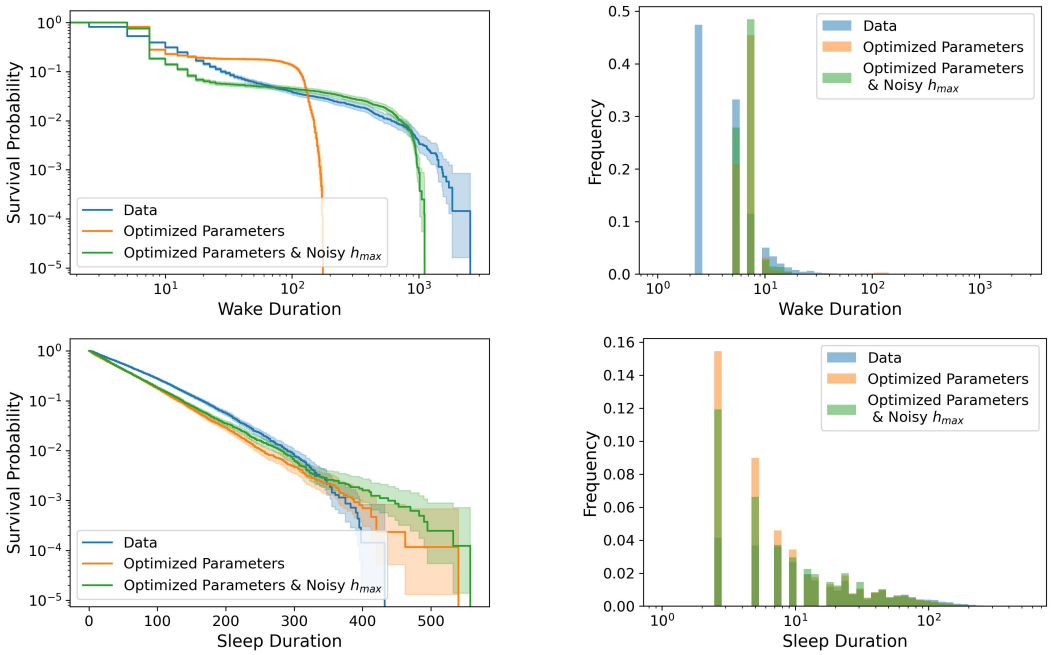


Figure 6.17: **Population-level similarities between data and model improve for optimized model with both noisy stimuli and noisy h_{max} input.** Data refers to the measures computed for the 100 samples of augmented light phase data. Optimized Parameters denote the model with (6.12), (6.13), (6.14) and noisy stimuli. Optimized Parameters & Noisy h_{max} denotes the fully optimized noisy model with appended noisy h_{max}.

6.5.3 Fine-tuning Weakly Correlated Parameters Enhances Model Fit

Finally, fixing r, τ_{hw} and τ_{hs} at their previously converged optimal values (see (6.12), (6.13)) and holding other parameters (including h_{max}) at their baseline values, we optimize a set of weakly influential yet biologically important parameters: β_W, α_W, α_S and $g_{G,W}$. The converged values for this parameter set are:

$$\beta_W = -0.80238 \pm 0.22721, \tag{6.14a}$$

$$\alpha_W = \quad 0.79447 \pm 0.03934, \tag{6.14b}$$

$$\alpha_S = \quad 0.21489 \pm 0.05928, \tag{6.14c}$$

$$g_{G,W} = \quad 5.99382 \pm 1.67424. \tag{6.14d}$$

The MMD value at the mean of this optimized parameter set is 0.1115, showing im-

provement over prior optimization stages. However, consistent with earlier findings, the base model with only noisy external stimuli does not sufficiently capture the distributional structure of the empirical data.

To address this, we append the noisy h_{max} component to the final set of optimized parameters. This addition leads to a substantial reduction in MMD to 0.1031, marking the best-performing configuration across all tested model variants.

As shown in Figure 6.16, the match between the data and model in terms of mean and standard error for various inference measures is considerably improved. Furthermore, Figure 6.17 reveals a closer fit between the data and model at the population level, particularly in the shape and spread of survival curves and duration histograms, with some shifted dynamics observed in the model outcomes (slightly longer wake and sleep durations).

Importantly, the Wasserstein distance between the model-generated and empirical distributions also reduces when noisy h_{max} is appended to the model with the final set of optimized parameters, see Table 6.3. Adding noisy h_{max} substantially changes the curvature and overall fit of the wake survival curves, yielding the lowest Wasserstein distance for this final model iteration. However, evidently, the curves or the distance do not vary much for the distribution of sleep durations. Arguably, when considering an overall fit for the data, the optimized set of parameters with both noisy stimuli and noisy h_{max} indicates a substantial improvement in capturing not only individual sample-level statistics but also the broader population-level dynamics.

6.6 Discussion

Mathematical modeling of irregular, polyphasic, and noisy sleep-wake data presents several challenges. In particular, estimating physiologically meaningful parameters from such data requires bridging gaps in three key areas: incorporating noise into otherwise mechanistic models, selecting appropriate distribution-level error metrics, and performing efficient optimization in high-dimensional parameter spaces. In this work, we address these challenges

by proposing a gradient-based framework for parameter estimation using rodent sleep-wake recordings.

Past works on modeling sleep-wake regulation often focus on fitting mean values or total sleep/wake durations, an approach that might overlook the finer temporal architectures of real-world data. We introduce a novel strategy using a flip-flop ODE model that captures transitions between sleep and wake states along with other characteristics. Treating each sleep-wake time trace as a sample from an underlying distribution, we shift the modeling objective toward matching distributions rather than point estimates. This perspective allows us to optimize parameters using a distributional distance metric, maximum mean discrepancy (MMD), which effectively captures fine-grained differences between model outputs and data.

We implement two algorithmic strategies for introducing noise into the flip-flop model: (1) additive noise stimuli affecting the firing rate of the wake-promoting population, and (2) a time-varying, noisy homeostatic threshold that modulates transitions between wake and sleep. The first noise mechanism models peripheral inputs into the wake-promoting populations and fragments the deterministic dynamics, producing short wake bouts that mirror rodent behavior during the light phase. The second introduces variability around the homeostatic transition threshold, producing longer wake bouts and adding distributional changes for wake durations without disrupting sleep durations. Crucially, both forms of noise are incorporated in a way that preserves differentiability, enabling the use of automatic differentiation, a core contribution of our framework.

Our results reveal that both noise components are essential for reproducing the empirical sleep-wake patterns. The noise stimulus alone closely approximates sleep durations but under-represents medium-length wake bouts (10–500 seconds) in rodents. In contrast, adding noise to the homeostatic threshold successfully introduces these durations without substantially disturbing sleep dynamics. This emphasizes that incorporating variability across biological processes is crucial for accurately modeling sleep-wake behavior, especially when accounting for the diversity seen across age and disease conditions.

For the current setting, when both components are combined, our model achieves its best performance across multiple levels of evaluation. We demonstrate the sample-level variability using MMD and descriptive inference statistics (e.g., mean and total durations) and population-level structure via survival curves, histograms, and Wasserstein distances.

Beyond modeling accuracy, this work contributes methodological tools for effective parameter estimation, including sensitivity analysis, structured optimization, and model evaluation through multiple distance metrics. Looking forward, several directions could further improve the model. Numerical solvers with finer temporal resolution, better handling of short-duration transitions, and fine-tuned optimization hyperparameters may enhance performance. While the flip-flop model offers interpretability and tractability, it is also limited in scope. Incorporating additional neuronal populations, such as REM-promoting neurons or other wake-promoting regions involved in sustaining longer wake bouts, could lead to more comprehensive and biologically realistic models.

Importantly, recent studies suggest that REM sleep disorders are early biomarkers of neurodegenerative diseases such as Parkinson's and dementia [142]. Extending our current framework to explore variability in physiological parameters affecting REM dynamics and neuronal excitability could therefore enhance its applicability to clinical research.

CHAPTER 7

Conclusions and Future Work

In this dissertation, we employed a combination of analytical, numerical and stochastic tools to study nonlinear dynamics arising in complex biological systems. In Chapters 2 through 5, we analyzed the validity and accuracy of reduced models characterized by quasi-steady-state (QSS) dynamics in enzymatic chemical reactions. In Chapter 6, we developed a data-driven optimization framework to infer physiological parameters that regulate sleep-wake behavior in rodents. Our work contributes to the overall understanding of domain-specific and simplistic models, examining their utility in capturing the true behavior of intricate biological processes.

In particular, QSS approximations yield reduced and tractable models that explicate the slow dynamics of two-timescale chemical systems. These are particularly relevant for applications in pharmacology and drug development, where interpretability and reliability are essential. To that end, we rigorously analyzed enzymatic mechanisms and their well-known QSS approximations, focusing on the Michaelis–Menten (MM) and Intermolecular Autocatalytic Zymogen Activation (IAZA) mechanisms. Our broader aim is to extend these mathematical frameworks to more complex systems both within and beyond chemical kinetics.

In Chapter 2, we derive QSS approximations for systems in the stochastic regime, reflecting the noise inherent in biochemical reactions. We revisit classical validity conditions for the MM approximation through both stochastic (Chapter 2) and deterministic (Chapter 3) lenses, proposing improved qualifying conditions and tighter error estimates. Using standard

tools of phase-plane analysis, we efficiently locate the slow manifold and ensure that the most optimal approximation is recognized. Extending this to more complex enzymatic systems is an interesting area of future work.

It is important to note that our analysis primarily focused on the approximation error in the slow regime, specifically the error between the QSS reduction and the actual flow on the slow manifold. However, two primary sources of error contribute to the overall accuracy of approximate solutions to singularly perturbed ODEs. The first is the error in approximating the flow on the slow manifold itself, as addressed in our analyses. The second is the error in approximating the trajectory's approach to the slow manifold, demarcating the initial transient (fast) dynamics. This latter source of error is still substantial when dealing with approximations but generally much more challenging to analyze. Although some progress has been made in this area (see [51]), further exploration and improvement is required.

Our work in Chapters 4 and 5 addresses key areas where classical reduction theory faces limitations. When systems exhibit singularities such as bifurcations, the assumptions required for direct application of reduction theory are not met. Our work analyses the IAZA model under an emerging transcritical bifurcation, providing alternative frameworks that validate QSS approximations in such settings. Furthermore, classical reduction theory doesn't favor systems that are not present in a "standard singular perturbation form", a gap we address using differential equation analysis. Our methods are generalizable and can be extended to analyze more involved models that include higher-order reactions and cooperative binding. A promising future direction is to systematically apply these techniques to biological systems with rich nonlinear features such as canards, folds, and chaotic regimes, thereby improving the reliability of reduced-order modeling in complex settings.

Blow-up techniques, traditionally developed for ODE systems, have also been extended to the study of stochastic models with stochastic differential equations (SDEs) and spatial models such as reaction-diffusion partial differential equations (PDEs) [87, 82]. In these settings, blow-up analysis is often used to resolve singularities and characterize their spatial

localization and structure. An exciting implementation of the blow-up method is to investigate whether critical manifolds and bifurcation structures in stochastic or spatial enzymatic models exhibit analogous non-hyperbolic features. This would not only extend the mathematical toolkit for analyzing singular perturbations in infinite dimensions but also deepen our understanding of spatial and stochastic effects in biochemical regulation.

In Chapter 6, we present an optimization pipeline to estimate sleep-wake regulatory parameters to capture the polyphasic and noisy dynamics of rodent sleep data. The novelty of the approach lies in using gradient-based optimization for a noisy ordinary differential equation (ODE) model where we directly compare model-generated time traces with observed rodent sleep score data, leveraging distributional distances to evaluate their agreement. We propose effective sample-level and population-level metrics to evaluate optimization outcomes, and importantly, our framework requires no preprocessing or modification of discrete data. This makes it adaptable for use in other domains where time-series data are common, such as behavioral science, neuroscience, or systems biology.

In this work, we implement a simple two-state flip-flop model that offers computational tractability while laying the foundation for more complex models. Future work can extend this framework to three-state models that distinguish between Rapid Eye Movement (REM) and Non-REM sleep as well, an essential distinction in sleep studies. Additionally, our analyses were restricted to light-phase data, during which rodents are generally less active. Extending the framework to dark-phase data, characterized by longer wake bouts and heightened activity, would offer a more complete understanding of rodent sleep architecture. A natural extension is to incorporate circadian inputs into the model and examine how regulatory parameters evolve under different lighting and behavioral conditions.

Notably, including the suprachiasmatic nucleus (SCN) may help elucidate the behavioral differences between the light and dark phases. The suprachiasmatic nucleus (SCN) functions as the master circadian pacemaker, critically controlling the timing and consolidation of the sleep-wake cycle with the light-dark cycle [122, 40]. Collectively, these mechanisms

have been well-studied and formalized in various mathematical models across mammalian species and have been shown to explain interspecies variation in sleep timing and structure [166, 146, 138]. These extensions would make the modeling framework both more general and more physiologically informative.

Beyond rodent studies, the presented framework might be scaled to human sleep research as well. With the growing availability of wearable devices that track sleep and activity patterns through actigraphy and physiological signals, there is a pressing need for models that can accommodate real-world, noisy, and irregularly sampled data. Since our optimization pipeline operates directly on unmodified discrete time traces, it is naturally suited for such settings. Future work could involve adapting the model to human-specific sleep features including longer circadian cycles, variable sleep architecture, and individualized sleep homeostasis through noisy processes. This would expand the framework's utility from controlled lab settings to population-scale, personalized sleep health assessments.

APPENDIX A

Additional Information and Calculations for Chapter 3

This appendix chapter provides supplementary information to enhance the understanding of the main text of Chapter 3.

A.1 Singular perturbation analysis of large k_{-1}

In perturbation form, the Michaelis-Menten mass action equations with large k_{-1} is

$$\dot{s} = -k_1(e_0 - c)s + \varepsilon^{-1}\widehat{k}_{-1}c, \tag{A.1a}$$

$$\dot{c} = \quad k_1(e_0 - c)s - k_2 c - \varepsilon^{-1}\widehat{k}_{-1}c, \tag{A.1b}$$

which admits the corresponding layer problem:

$$\dot{s} = \varepsilon^{-1}\widehat{k}_{-1}c, \tag{A.2a}$$

$$\dot{c} = -\varepsilon^{-1}\widehat{k}_{-1}c. \tag{A.2b}$$

The critical manifold is therefore s-axis and is normally hyperbolic and attracting. The projection matrix, π^s, is given by

$$\pi^s = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}. \tag{A.3}$$

The Fenichel reduction is therefore

$$\begin{pmatrix} \dot{s} \\ \dot{c} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -k_1(e_0 - c)s + \varepsilon^{-1}\widehat{k}_{-1}c \\ k_1(e_0 - c)s - k_2 c - \varepsilon^{-1}\widehat{k}_{-1}c \end{pmatrix} \Bigg|_{c=0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{A.4}$$

which is trivial. Thus, one must appeal to higher-order terms in order to recover a non-trivial QSSA.

A.2 QSS approximations and initial conditions

Here we briefly discuss how to compute initial conditions for quasi-steady-state approximations. For planar systems with a normally hyperbolic and attracting critical manifold, S_0, the two-dimensional stable manifold of S_0, denoted as $W^s(S_0)$, is comprised of one-dimensional manifolds called *fast* fibers. Each point $p_{\mathcal{B}} \in S_0$ is called a base point, and each base point is a steady-state (equilibrium) solution to the layer problem. Furthermore, there is a fiber, $\mathcal{F}_0(p_{\mathcal{B}})$, corresponding to every base point belonging to S_0, which is simply the stable manifold of the base point, $p_{\mathcal{B}}$:

$$\mathcal{F}_0(p_{\mathcal{B}}) =: W^s(p_{\mathcal{B}}), \quad W^s(S_0) = \bigcup_{p_{\mathcal{B}} \in S_0} \mathcal{F}_0(p_{\mathcal{B}}).$$

The base points themselves serve as approximate initial conditions for quasi-steady-state reductions. Thus, the task is to identify the unique base point a given trajectory approaches during the transient (fast) phase of the reaction, and then equip the QSS reduction with this base point as its initial condition.

As a preliminary example, let's consider the Michaelis-Menten mass action system with small k_2. The perturbation form of the mass action system is

$$\dot{s} = -k_1(e_0 - c)s + k_{-1}c \tag{A.5a}$$

$$\dot{c} = \quad k_1(e_0 - c)s - k_{-1}c - \varepsilon\widehat{k}_2 c, \tag{A.5b}$$

and so the corresponding layer problem is

$$\dot{s} = -k_1(e_0 - c)s + k_{-1}c, \tag{A.6a}$$

$$\dot{c} = \quad k_1(e_0 - c)s - k_{-1}c. \tag{A.6b}$$

The critical manifold is

$$S_0 := \left\{ (s, c) \in \mathbb{R}^2_{\geq 0} : c = \frac{e_0 s}{K_S + s} \right\} \tag{A.7}$$

and therefore each base point, $p_{\mathcal{B}}$, has the representation

$$p_{\mathcal{B}} = (s, e_0 s/(K_S + s)), \quad \forall s \in [0, s_0]. \tag{A.8}$$

The typical initial condition for the Michaelis-Menten reaction network is $(s, c)(0) = (s_0, 0)$ and, as previously mentioned, the Fenichel reduction for small k_2 is

$$\begin{pmatrix} \dot{s} \\ \dot{c} \end{pmatrix} = -\frac{k_2 e_0 s(K_S + s)}{K_S e_0 + (K_S + s)^2} \begin{pmatrix} 1 \\ \dfrac{K_S e_0 - K(K_S + s)}{(K_S + s)^2} \end{pmatrix}. \tag{A.9}$$

The question is whether or not we can equip (A.9) with the same initial condition, $(s, c)(0) = (s_0, 0)$ and expect the solution to the mass action system to converge to the solutions to (A.9) as $k_2 \to 0$?

The answer to this question is no: The reduction (A.9) approximates the dynamics *on* the slow manifold; it does not describe the dynamics that emerge in the approach to the slow manifold. Clearly, solutions to the mass action equations (A.5) capture the fast *and* slow stages of the reaction, but the QSS reduction (A.9) describes only the slow stage. Thus, we need to equip (A.9) with a "modified" initial condition, $\tilde{x}_0 = (\tilde{s}, \tilde{c})$, but how do we obtain $\tilde{x}_0$? The short answer is that the modified initial condition is the associated base point of the fast fiber that contains $(s_0, 0)$. In this example, the fiber that contains $(s, c)(0) = (s_0, 0)$

is

$$\mathcal{F}_0(p_{\mathcal{B}}) = \{(s,c) \in \mathbb{R}^2_{\geq 0} : s + c = s_0\} \tag{A.10}$$

and the modified initial condition is simply the intersection of $\mathcal{F}_0(p_{\mathcal{B}})$ with S_0:

$$\tilde{x}_0 = (s, e_0 s/(K_S + s)), \quad \text{such that } s \text{ satisfies: } \; s + \frac{e_0 s}{K_S + s} = s_0, \tag{A.11}$$

which can be solved algebraically using the quadratic formula:

$$\tilde{s} = \frac{1}{2}\left(s_0 - e_0 - K_S + \sqrt{(s_0 - e_0 - K_S)^2 + s_0 K_S}\right), \quad \tilde{c} = e_0 \tilde{s}/(K_S + \tilde{s}). \tag{A.12}$$

The takeaway from this example is that if we equip the reduction (A.9) with the modified initial condition $(s,c)(0) = (\tilde{s}, \tilde{c})$, then the solution to this system will converge to the solution of the mass action system (A.5) (also equipped with the initial condition $(s,c)(0) = (\tilde{s}, \tilde{c})$) as $k_2 \to 0$.

Next consider the case for small e_0, which is actually more difficult than small k_2. As already mentioned, the critical manifold is the s-axis, but the QSS variety is actually the c-nullcline, and in fact this distinction was not fully understood until recently; see Noethen and Walcher [127] for details. The base points in this case are $p_{\mathcal{B}} = (s, 0)$, but the fibers associated with each base point are rather complicated (again, see Noethen and Walcher [127], Section 3.1 for a discussion concerning the first integral of the layer problem associated with small e_0). However, since $(s_0, 0) \in S_0$, we can equip the MM rate law,

$$\dot{s} = -\frac{k_2 e_0 s}{K_M + s} \tag{A.13}$$

with the initial condition $s(0) = s_0$.

Initial conditions for the complex, c, can also be computed by analyzing the transient dynamics, but we will not go into those details here. Interested readers should consult Segel

and Slemrod [163] as well as Lax and Walcher [109].

APPENDIX B

Additional Information and Calculations for Chapter 4

These appendices provide supplementary information for Chapter 4 to support the main text.

B.1 Quasi-Homogeneity

Scaling $(\widehat{w}, \widehat{c}, \varepsilon)$ as $(r^{a_1}\widehat{w}, r^{a_2}\widehat{c}, r^{a_3}\varepsilon)$ and writing the conditions for quasi-homogeneity for the shifted system (4.46), we obtain

$$(r^{a_3}\varepsilon)k_2(r^{a_2}\widehat{c} + E_T/2) = r^{k+a_1}F_1(\widehat{w}, \widehat{c}, \varepsilon),$$
$$k_1((r^{a_2}\widehat{c})^2 - (r^{a_1}\widehat{w})^2) - (r^{a_3}\varepsilon)(k_{-1} + k_2)(r^{a_2}\widehat{c} + E_T/2) = r^{k+a_2}F_2(\widehat{w}, \widehat{c}, \varepsilon).$$

To solve for the type and degree of quasi-homogeneity, we compare the degree of r on both sides. Similar to the normal form analysis in Krupa and Szmolyan [103], we can ignore the term of $\mathcal{O}(\varepsilon c)$ in both the functions as a higher order term in r for this purpose. Solving $a_3 = k + a_1$ and $2a_1 = 2a_2 = a_3 = k + a_2$; we obtain $a_1 = 1, a_2 = 1, a_3 = 2$ and $k = 1$ as possible values. Thus, the system is quasi-homogeneous of the type (1,1,2) and degree 2. Thus, the analysis in Section 4.5.2 follow.

B.2 Coordinate Charts for a Blow-Up Analysis

For the IAZA system in $(\widehat{w}, \widehat{c})$ coordinates, we define regions of the phase plane where

the solution trajectories enter and exit.

Let $\rho > 0$ be a small constant and $I \in \mathbb{R}$ be a small interval around 0. We define

$$\Delta^{in} = \{(-\rho, \widehat{c}), \widehat{c} + \rho \in I\}, \tag{B.1}$$

$$\Delta^{out} = \{(\rho, \widehat{c}), \widehat{c} + \rho \in I\} \tag{B.2}$$

to demarcate the left half and right half of the plane. This suggests that we are most interested in the regions $B^-_{\bar{w}} := B \cap \{\bar{w} < 0\}$ and $B^+_{\bar{w}} := B \cap \{\bar{w} > 0\}$. Thus, we define local rectangular coordinates in these regions to make our computation easier. This is done through intermediary directional blow-ups in the directions of $\bar{w}$, $\bar{\varepsilon}$ and $-\bar{w}$ (more details shortly). We need three charts to describe all the dynamics: K_1, K_2 and K_3:

The chart K_1 is defined by

$$\widehat{w} = -r_1 \qquad \widehat{c} = r_1 c_1 \qquad \varepsilon = r_1^2 \varepsilon_1, \tag{B.3}$$

the rescaling chart K_2 is defined by

$$\widehat{w} = r_2 w_2 \qquad \widehat{c} = r_2 c_2 \qquad \varepsilon = r_2^2, \tag{B.4}$$

and finally chart K_3 is defined by

$$\widehat{w} = r_3 \qquad \widehat{c} = r_3 c_3 \qquad \varepsilon = r_3^2 \varepsilon_3. \tag{B.5}$$

Intuitively, we can think of chart K_1 covering parts of B where trajectories enter and the chart K_3 covering parts where trajectories exit. The chart K_2 is known as the rescaling chart and covers the blown-up neighborhood of $(0, 0)$. This is because r_2 is fixed ($\dot{\varepsilon} = 0$) and the coordinate transformation is essentially a rescaling of $\widehat{w}$ and $\widehat{c}$. The steps followed to obtain the chart K_2 are shown in detail below.

Recall equation (4.47), the polar quasi-homogeneous blow-up is given by $\phi : S^2 \times [0, r_0] \to \mathbb{R}^3$. In order to take the directional blow-up in the ε-direction, we introduce $\kappa_2 : B \to \mathbb{R}^3$ such that

$$\kappa_2(\bar{w}, \bar{c}, \bar{\varepsilon}, r) := \left(\frac{\bar{w}}{\bar{\varepsilon}^{b_1}}, \frac{\bar{c}}{\bar{\varepsilon}^{b_2}}, \frac{r}{\bar{\varepsilon}^{b_3}} \right) = (w_2, c_2, r_2). \tag{B.6}$$

The idea is that we should be able to eliminate the variable $\bar{\varepsilon}$ in the space away from $\{\bar{\varepsilon} = 0\}$. The blown-up map on $\mathbb{R}^3$ extends to the blown-up point, $\{\bar{\varepsilon} = 0\}$, using the same techniques as the extension of a polar blow-up to $\{r = 0\}$ (refer [106], theorem 7.1.6). Here, (w_2, c_2, r_2) are the local rectangular coordinates. The coefficients b_i are determined such that $\phi = \mu_2 \circ \kappa_2$ where μ_2 is the directional blow-up in terms of the local coordinates.

From ϕ, we have $\widehat{w} = r\bar{w}, \widehat{c} = r\bar{c}$ and $\varepsilon = r^2\bar{\varepsilon}$.

From κ_2, we have $\bar{w} = \bar{\varepsilon}^{b_1} w_2, \bar{c} = \bar{\varepsilon}^{b_2} c_2$, and $r = \bar{\varepsilon}^{b_3} r_2$.

We define $\mu_2 : \mathbb{R}^3 \to \mathbb{R}^3$ such that

$$\begin{aligned}
\widehat{w} = r\bar{w} &= \bar{\varepsilon}^{b_3} r_2 \bar{\varepsilon}^{b_1} w_2, \\
\widehat{c} = r\bar{c} &= \bar{\varepsilon}^{b_3} r_2 \bar{\varepsilon}^{b_2} c_2, \\
\varepsilon = r^2\bar{\varepsilon} &= (\bar{\varepsilon}^{b_3} r_2)^2 \bar{\varepsilon}.
\end{aligned}$$

To remove $\bar{\varepsilon}$, $b_1 = 1/2, b_2 = 1/2$ and $b_3 = -1/2$. Finally, for these values, the chart K_2 is defined by (4.48). The chart K_1 and K_3 are computed by introducing a directional blow-up in the direction of $\bar{w}$ and $-\bar{w}$ respectively.

APPENDIX C

Additional Information for Chapter 6

This appendix chapter provides supplementary information to enhance the understanding of the main text of Chapter 6.

C.1 Inference Measures for Experimental Data

In this section, we provide the details of the experimental dataset used in this chapter. The tables include the inference measures observed for different mice samples. The light phase data has 50 recordings divided into two 2-hour blocks: 9 AM to 11 AM and 11 AM to 1 PM. The dark phase data has 48 recordings observed over 9 PM to 11 PM.

Table for the descriptive statistics of light phase data in time block 1 - 9 AM to 11 AM

Mouse ID	Mean Wake Duration	Mean Sleep Duration	Number of Wake Bouts	Number of Sleep Bouts	Total Wake Duration	Total Sleep Duration
c18_020618n1	26.7	66.8	77.0	77.0	2057.5	5142.5
c19_020418n1	17.5	86.8	69.0	69.0	1210.0	5990.0
c19_020518n1	28.9	75.5	69.0	69.0	1992.5	5207.5
c20_020418n1	29.4	77.6	68.0	67.0	2000.0	5200.0
c20_020618n1	48.8	72.0	59.0	60.0	2877.5	4322.5

c21_020418n1	18.2	85.0	69.0	70.0	1252.5	5947.5
c21_020518n1	18.1	93.0	64.0	65.0	1157.5	6042.5
c21_020618n1	21.3	86.5	66.0	67.0	1405.0	5795.0
c27_030318n1	26.5	76.4	70.0	70.0	1852.5	5347.5
c27_030418n1	21.3	77.6	72.0	73.0	1532.5	5667.5
c28_030218n1	19.8	67.2	82.0	83.0	1625.0	5575.0
c28_030318n1	11.4	85.9	74.0	74.0	845.0	6355.0
c28_030418n1	34.4	92.5	56.0	57.0	1927.5	5272.5
c29_030218n1	27.0	74.4	71.0	71.0	1920.0	5280.0
c29_030318n1	21.4	63.6	84.0	85.0	1795.0	5405.0
c29_030418n1	17.2	55.0	99.0	100.0	1705.0	5495.0
c30_030218n1	27.9	86.4	63.0	63.0	1757.5	5442.5
c30_030318n1	14.5	60.7	95.0	96.0	1375.0	5825.0
c30_030418n1	13.4	85.4	72.0	73.0	967.5	6232.5
c36_032418n1	31.5	87.1	60.0	61.0	1887.5	5312.5
c36_032518n1	20.5	68.4	81.0	81.0	1657.5	5542.5
c36_032618n1	25.9	78.4	69.0	69.0	1787.5	5412.5
c37_032418n1	31.6	80.9	64.0	64.0	2022.5	5177.5
c37_032518n1	26.6	91.8	60.0	61.0	1597.5	5602.5
c37_032618n1	32.7	76.9	65.0	66.0	2125.0	5075.0
c38_032418n1	22.8	124.6	48.0	49.0	1092.5	6107.5
c38_032518n1	38.9	114.3	47.0	47.0	1827.5	5372.5
c38_032618n1	61.5	67.1	56.0	56.0	3445.0	3755.0
c42_040918n1	19.0	52.3	101.0	101.0	1922.5	5277.5
c42_041018n1	18.7	73.8	77.0	78.0	1440.0	5760.0
c42_041118n1	43.9	72.3	62.0	62.0	2720.0	4480.0
c43_040918n1	39.5	79.8	61.0	60.0	2412.5	4787.5

c43_041018n1	33.0	68.4	71.0	71.0	2342.5	4857.5
c43_041118n1	28.0	73.8	70.0	71.0	1960.0	5240.0
c44_041018n1	31.9	74.0	68.0	68.0	2170.0	5030.0
c44_041118n1	30.2	71.6	70.0	71.0	2117.5	5082.5
c45_040918n1	17.8	102.2	60.0	60.0	1070.0	6130.0
c45_041018n1	18.7	81.6	71.0	72.0	1327.5	5872.5
c45_041118n1	44.7	82.4	56.0	57.0	2505.0	4695.0
c47_042218n1	30.6	86.9	62.0	61.0	1900.0	5300.0
c47_042318n1	32.7	89.9	58.0	59.0	1895.0	5305.0
c48_042218n1	11.0	92.0	69.0	70.0	757.5	6442.5
c48_042318n1	77.4	104.5	39.0	40.0	3020.0	4180.0
c48_042418n1	46.2	70.6	61.0	62.0	2820.0	4380.0
c49_042218n1	35.6	81.1	61.0	62.0	2172.5	5027.5
c49_042318n1	27.1	88.6	63.0	62.0	1707.5	5492.5
c49_042418n1	27.0	76.2	69.0	70.0	1865.0	5335.0
c50_042218n1	31.8	68.6	71.0	72.0	2257.5	4942.5
c50_042318n1	33.9	84.7	60.0	61.0	2035.0	5165.0
c50_042418n1	22.9	80.2	69.0	70.0	1582.5	5617.5

Table for the descriptive statistics of the dark phase data in time block 9 PM to 11 PM

Mouse ID	Mean Wake Duration	Mean Sleep Duration	Number of Wake Bouts	Number of Sleep Bouts	Total Wake Duration	Total Sleep Duration
11_022219n1	377.3	72.7	16.0	16.0	6037.5	1162.5
12_022219n1	137.2	83.5	33.0	32.0	4527.5	2672.5

14_022219n1	186.3	59.9	29.0	30.0	5402.5	1797.5
14_022419n1	316.6	65.8	19.0	18.0	6015.0	1185.0
9_022219n1	109.2	60.8	42.0	43.0	4585.0	2615.0
9_022419n1	166.8	63.4	31.0	32.0	5170.0	2030.0
HA32_081618n1	216.8	83.2	24.0	24.0	5202.5	1997.5
HA37_081618n1	107.5	36.5	50.0	50.0	5372.5	1827.5
HA38_081618n1	94.1	55.9	48.0	48.0	4515.0	2685.0
HA40_081618n1	116.7	39.8	46.0	46.0	5367.5	1832.5
HA41_081618n1	112.6	47.4	45.0	45.0	5067.5	2132.5
HA42_081618n1	139.4	40.6	40.0	40.0	5575.0	1625.0
HA43_081618n1	234.9	43.7	26.0	25.0	6107.5	1092.5
HA44_081618n1	153.6	36.8	38.0	37.0	5837.5	1362.5
HA52_090718n1	125.4	24.6	48.0	48.0	6017.5	1182.5
HA55_090718n1	1380.0	60.0	5.0	5.0	6900.0	300.0
HA57_090718n1	542.3	57.7	12.0	12.0	6507.5	692.5
HA58_090718n1	79.8	31.5	65.0	64.0	5185.0	2015.0
HA59_090718n1	191.7	41.9	31.0	30.0	5942.5	1257.5
SD32_042419n1	82.5	23.7	68.0	67.0	5610.0	1590.0
SD32_042619n1	128.9	65.7	37.0	37.0	4770.0	2430.0
SD32_042819n1	101.8	71.3	42.0	41.0	4275.0	2925.0
SD33_042619n1	195.4	36.9	31.0	31.0	6057.5	1142.5
SD33_042819n1	331.8	72.2	18.0	17.0	5972.5	1227.5
SD34_042419n1	568.1	31.9	12.0	12.0	6817.5	382.5
SD34_042619n1	46.3	33.7	90.0	90.0	4167.5	3032.5
SD34_042819n1	80.9	66.0	49.0	49.0	3965.0	3235.0
SD35_042419n1	221.6	27.7	29.0	28.0	6425.0	775.0
SD35_042619n1	216.3	63.0	26.0	25.0	5625.0	1575.0

SD35_042819n1	283.9	76.1	20.0	20.0	5677.5	1522.5
SD36_042419n1	273.7	111.1	19.0	18.0	5200.0	2000.0
SD36_042819n1	586.7	13.3	12.0	12.0	7040.0	160.0
SD37_042419n1	426.2	57.7	15.0	14.0	6392.5	807.5
SD37_042819n1	2390.0	15.0	3.0	2.0	7170.0	30.0
SD38_042419n1	146.5	53.5	36.0	36.0	5272.5	1927.5
SD38_042819n1	55.1	24.7	90.0	91.0	4955.0	2245.0
SD39_042419n1	228.6	88.3	23.0	22.0	5257.5	1942.5
SD40_042519n1	609.1	50.0	11.0	10.0	6700.0	500.0
SD40_042919n1	275.5	67.4	21.0	21.0	5785.0	1415.0
SD41_042919n1	213.4	66.1	26.0	25.0	5547.5	1652.5
SD42_042919n1	122.3	72.3	37.0	37.0	4525.0	2675.0
SD44_042519n1	64.2	47.3	65.0	64.0	4172.5	3027.5
SD44_042919n1	163.9	76.1	30.0	30.0	4917.5	2282.5
SD45_042519n1	573.3	29.1	12.0	11.0	6880.0	320.0
SD45_042919n1	136.0	75.7	34.0	34.0	4625.0	2575.0
SD46_042519n1	185.4	71.7	28.0	28.0	5192.5	2007.5
SD46_042919n1	192.9	57.3	29.0	28.0	5595.0	1605.0
SD47_042919n1	70.4	88.0	46.0	45.0	3240.0	3960.0

Table for the descriptive statistics of the light phase data in time block 2 - 11 AM to 1 PM

Mouse ID	Mean Wake Duration	Mean Sleep Duration	Number of Wake Bouts	Number of Sleep Bouts	Total Wake Duration	Total Sleep Duration
c18_020618n1	25.5	69.5	75.0	76.0	1915.0	5285.0

c19_020418n1	44.8	65.2	66.0	65.0	2960.0	4240.0
c19_020518n1	35.0	46.8	88.0	88.0	3077.5	4122.5
c20_020418n1	36.8	55.5	78.0	78.0	2867.5	4332.5
c20_020618n1	15.6	67.3	86.0	87.0	1342.5	5857.5
c21_020418n1	20.7	59.5	89.0	90.0	1845.0	5355.0
c21_020518n1	26.3	81.6	66.0	67.0	1735.0	5465.0
c21_020618n1	26.1	81.8	66.0	67.0	1722.5	5477.5
c27_030318n1	30.6	71.3	70.0	71.0	2140.0	5060.0
c27_030418n1	29.6	73.6	69.0	70.0	2045.0	5155.0
c28_030218n1	13.6	55.1	104.0	105.0	1412.5	5787.5
c28_030318n1	23.1	70.5	77.0	77.0	1775.0	5425.0
c28_030418n1	40.9	75.9	61.0	62.0	2492.5	4707.5
c29_030218n1	28.4	59.5	82.0	82.0	2325.0	4875.0
c29_030318n1	19.0	61.9	89.0	89.0	1695.0	5505.0
c29_030418n1	34.2	56.2	79.0	80.0	2700.0	4500.0
c30_030218n1	22.0	72.8	76.0	76.0	1670.0	5530.0
c30_030318n1	22.3	60.5	87.0	87.0	1940.0	5260.0
c30_030418n1	16.1	60.7	93.0	94.0	1495.0	5705.0
c36_032418n1	19.2	86.9	67.0	68.0	1287.5	5912.5
c36_032518n1	23.9	57.0	89.0	89.0	2125.0	5075.0
c36_032618n1	25.2	59.5	85.0	85.0	2140.0	5060.0
c37_032418n1	23.7	77.7	71.0	71.0	1685.0	5515.0
c37_032518n1	23.2	81.5	68.0	69.0	1577.5	5622.5
c37_032618n1	24.0	94.1	61.0	61.0	1462.5	5737.5
c38_032418n1	23.2	98.8	59.0	59.0	1370.0	5830.0
c38_032518n1	34.8	112.1	49.0	49.0	1705.0	5495.0
c38_032618n1	15.2	77.3	77.0	78.0	1167.5	6032.5

c42_040918n1	19.6	61.3	89.0	89.0	1747.5	5452.5
c42_041018n1	79.5	67.4	49.0	49.0	3897.5	3302.5
c42_041118n1	30.1	55.6	84.0	84.0	2527.5	4672.5
c43_040918n1	30.6	75.3	68.0	68.0	2082.5	5117.5
c43_041018n1	34.9	60.3	75.0	76.0	2617.5	4582.5
c43_041118n1	34.1	73.3	67.0	67.0	2287.5	4912.5
c44_041018n1	29.3	73.1	71.0	70.0	2082.5	5117.5
c44_041118n1	25.2	72.4	73.0	74.0	1840.0	5360.0
c45_040918n1	25.2	80.7	68.0	68.0	1715.0	5485.0
c45_041018n1	20.7	64.3	84.0	85.0	1735.0	5465.0
c45_041118n1	33.2	85.4	60.0	61.0	1990.0	5210.0
c47_042218n1	30.7	104.6	54.0	53.0	1657.5	5542.5
c47_042318n1	77.8	87.6	43.0	44.0	3345.0	3855.0
c48_042218n1	80.1	89.2	42.0	43.0	3362.5	3837.5
c48_042318n1	33.5	103.0	52.0	53.0	1742.5	5457.5
c48_042418n1	53.0	76.5	55.0	56.0	2915.0	4285.0
c49_042218n1	28.1	90.4	60.0	61.0	1687.5	5512.5
c49_042318n1	37.2	73.6	65.0	65.0	2417.5	4782.5
c49_042418n1	33.8	75.8	65.0	66.0	2197.5	5002.5
c50_042218n1	51.3	77.3	56.0	56.0	2872.5	4327.5
c50_042318n1	34.1	72.2	67.0	68.0	2287.5	4912.5
c50_042418n1	36.4	82.2	60.0	61.0	2185.0	5015.0

C.2 Theoretical Distributions for the Sleep Data

In this section, we provide evidence that the inter-wake intervals roughly follow an exponential distribution. We use the 'distfit' package in Python to evaluate the goodness of fit

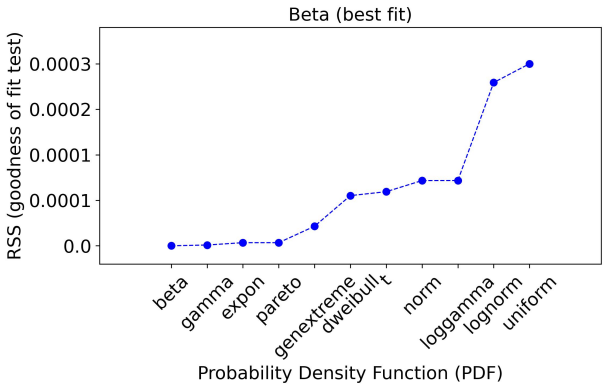


Figure C.1: **Exponential distribution provides good fit for the distribution of sleep durations in light phase data.** Other distributions such as beta, gamma and pareto also provide good fit. We use the python package 'distfit' for the plot.

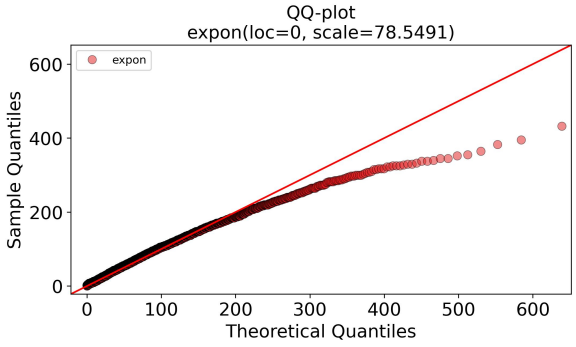


Figure C.2: **The Quantile-Quantile (QQ) plot compares the empirical distribution of sleep durations with the best fit theoretical exponential distribution.** It shows a good fit for the shorter durations of sleep but deviates from the 45-degree line. We use the python package 'distfit' for the plot.

for different theoretical distributions and find that an exponential distribution performs at par as some other light-tailed distributions as well, see Figure C.1.

The Quantile-Quantile-plot (QQ-plot), Figure C.2, depicts a good fit of the theoretical exponential distribution for the shorter durations. However, the sample quantiles deviate lower than the theoretical quantiles.

BIBLIOGRAPHY

[1] Milton Abramowitz, Irene A Stegun, and Robert H Romer. Handbook of mathematical functions with formulas, graphs, and mathematical tables, 1988.

[2] Peter Achermann and Alexander A Borbély. Mathematical models of sleep regulation. *Frontiers in Bioscience*, 8(6):1064, 2003.

[3] Peter Achermann, Derk-Jan Dijk, Daniel P Brunner, and Alexander A Borbély. A model of human sleep homeostasis based on EEG slow-wave activity: quantitative comparison of data and simulations. *Brain Research Bulletin*, 31(1-2):97–113, 1993.

[4] Ravi Allada, Chiara Cirelli, and Amita Sehgal. Molecular mechanisms of sleep homeostasis in flies and mammals. *Cold Spring Harbor Perspectives in Biology*, 9(8):a027730, 2017.

[5] Asohan Amarasingham, Ting-Li Chen, Stuart Geman, Matthew T Harrison, and David L Sheinberg. Spike count reliability and the Poisson hypothesis. *Journal of Neuroscience*, 26(3):801–809, 2006.

[6] Christina Athanasouli, Kelsey Kalmbach, Victoria Booth, and Cecilia G Diniz Behn. NREM–REM alternation complicates transitions from napping to non-napping behavior in a three-state model of sleep–wake regulation. *Mathematical Biosciences*, 355: 108929, 2023.

[7] Debashis Barik, Mark R. Paul, William T. Baumann, Yang Cao, and John J. Tyson. Stochastic simulation of enzyme-catalyzed reactions with disparate timescales. *Biophys. J.*, 95(8), 2008.

[8] Maxim Bazhenov, Igor Timofeev, Mircea Steriade, and Terrence J Sejnowski. Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. *Journal of Neuroscience*, 22(19):8691–8704, 2002.

[9] Cecilia G Diniz Behn and Victoria Booth. Modeling the temporal architecture of rat sleep-wake behavior. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4713–4716. IEEE, 2011.

[10] Matthew R Bennett, Dmitri Volfson, Lev Tsimring, and Jeff Hasty. Transient dynamics of genetic regulatory networks. *Biophys. J.*, 92(10):3501–3512, 2007.

[11] Nils Berglund and Barbara Gentz. *Noise-induced phenomena in slow-fast dynamical systems*. Springer-Verlag London, Ltd., London, 2006.

[12] Nils Berglund and Barbara Gentz. *Noise-induced phenomena in slow-fast dynamical systems: a sample-paths approach*. Probability and Its Applications. Springer Science & Business Media, 2006.

[13] Matt T Bianchi, Nathaniel A Eiseman, Sydney S Cash, Joseph Mietus, Chung-Kang Peng, and Robert J Thomas. Probabilistic sleep architecture models in patients with and without sleep apnea. *Journal of Sleep Research*, 21(3):330–341, 2012.

[14] Peter J Bickel and Kjell A Doksum. *Mathematical statistics: basic ideas and selected topics, volumes I-II package*. Texts in Statistical Science. Chapman and Hall/CRC, 2015.

[15] Terri Blackwell, Susan Redline, Sonia Ancoli-Israel, Jennifer L Schneider, Susan Surovec, Nathan L Johnson, Jane A Cauley, Katie L Stone, and Study of Osteoporotic Fractures Research Group. Comparison of sleep parameters from actigraphy and polysomnography in older women: the SOF study. *Sleep*, 31(2):283–291, 2008.

[16] Victoria Booth and Cecilia G Diniz Behn. Physiologically-based modeling of sleep–wake regulatory networks. *Mathematical Biosciences*, 250:54–68, 2014.

[17] Alexander A Borb and Peter Achermann. Sleep homeostasis and models of sleep regulation. *Journal of Biological Rhythms*, 14(6):559–570, 1999.

[18] Alexander A Borbély et al. A two process model of sleep regulation. *Human Neurobiology*, 1(3):195–204, 1982.

[19] David M Bortz, Daniel A Messenger, and Vanja Dukic. Direct estimation of parameters in ODE models using WENDy: Weak-form estimation of nonlinear dynamics. *Bulletin of Mathematical Biology*, 85(11):110, 2023.

[20] George Edward Briggs and John Burdon Sanderson Haldane. A note on the kinetics of enzyme action. *Biochemical Journal*, 19(2):338, 1925.

[21] Ritchie E Brown, Radhika Basheer, James T McKenna, Robert E Strecker, and Robert W McCarley. Control of sleep and wakefulness. *Physiological Reviews*, 2012.

[22] Matt S. Calder and David Siegel. Properties of the Michaelis-Menten mechanism in phase space. *J. Math. Anal. Appl.*, 339(2):1044–1064, 2008. ISSN 0022-247X,1096-0813.

[23] Nancy L Chamberlin, Elda Arrigoni, Thomas C Chou, Thomas E Scammell, Ross W Greene, and Clifford B Saper. Effects of adenosine on gabaergic synaptic inputs to identified ventrolateral preoptic neurons. *Neuroscience*, 119(4):913–918, 2003.

[24] Boseung Choi, Grzegorz A Rempala, and Jae Kyoung Kim. Beyond the Michaelis-Menten equation: Accurate and efficient estimation of enzyme kinetic parameters. *Scientific Reports*, 7(1):17018, 2017.

[25] Chiara Cirelli and Giulio Tononi. Is sleep essential? *PLoS Biology*, 6(8):e216, 2008.

[26] Matteo Colombo and Peggy Seriès. Bayes in the brain—On Bayesian Modelling in Neuroscience. *The British Journal for the Philosophy of Science*, 63:697–723, 2012.

[27] Geoffrey M Cooper. The central role of enzymes as biological catalysts. In *The Cell: A Molecular Approach. (2nd Edition)*. Sinauer Associates, 2000.

[28] Athel Cornish-Bowden. One hundred years of Michaelis–Menten kinetics. *Perspectives in Science*, 4:3–9, 2015.

[29] Michael J. Davis and Rex T. Skodje. Geometric investigation of low-dimensional manifolds in systems approaching equilibrium. *J. Chem. Phys.*, 111:859–874, 1999.

[30] Peter Dayan and Laurence F Abbott. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Computational Neuroscience. MIT press, 2005.

[31] Gustavo Deco, Viktor K Jirsa, Peter A Robinson, Michael Breakspear, and Karl Friston. The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Computational Biology*, 4(8):e1000092, 2008.

[32] Hartmut Derendorf and Bernd Meibohm. Modeling of pharmacokinetic/pharmacodynamic (PK/PD) relationships: concepts and perspectives. *Pharmaceutical Research*, 16:176–185, 1999.

[33] Derk-Jan Dijk and Charles A Czeisler. Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans. *Journal of Neuroscience*, 15(5):3526–3538, 1995.

[34] Cecilia G Diniz Behn and Victoria Booth. Simulating microinjection experiments in a novel model of the rat sleep-wake regulatory network. *Journal of Neurophysiology*, 103(4):1937–1953, 2010.

[35] Vivek Dua. An artificial neural network approximation based decomposition approach for parameter estimation of system of ordinary differential equations. *Computers & Chemical Engineering*, 35(3):545–553, 2011.

[36] P Duchêne and P Rouchon. Kinetic scheme reduction via geometric singular perturbation techniques. *Chemical Engineering Science*, 51(20):4661–4672, 1996.

[37] Freddy Dumortier. Singularities of vector fields on the plane. *Journal of Differential Equations*, 23(1):53–106, 1977.

[38] Freddy Dumortier and Robert H Roussarie. *Canard cycles and enter manifolds*, volume 121 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 1996.

[39] Justin R Dunmyre, George A Mashour, and Victoria Booth. Coupled flip-flop model for REM sleep regulation in the rat. *PLoS One*, 9(4):e94481, 2014.

[40] Amy Easton, Peter Meerlo, Bernard Bergmann, and Fred W Turek. The suprachiasmatic nucleus regulates sleep timing and amount in mice. *Sleep*, 27(7):1307–1318, 2004.

[41] Kristin Eckel-Mahan and Paolo Sassone-Corsi. Phenotyping circadian rhythms in mice. *Current Protocols in Mouse Biology*, 5(3):271–281, 2015.

[42] Justin Eilertsen and Santiago Schnell. The quasi-steady-state approximations revisited: Timescales, small parameters, singularities, and normal forms in enzyme kinetics. *Math. Biosci.*, 325:108339, 2020.

[43] Justin Eilertsen and Santiago Schnell. The quasi-steady-state approximations revisited: Timescales, small parameters, singularities, and normal forms in enzyme kinetics. *Mathematical Biosciences*, 325:108339, 2020.

[44] Justin Eilertsen, Wylie Stroberg, and Santiago Schnell. Characteristic, completion or matching timescales? An analysis of temporary boundaries in enzyme kinetics. *J. Theoret. Biol.*, 481:28–43, 2019. ISSN 0022-5193,1095-8541.

[45] Justin Eilertsen, Marc Roussel, Santiago Schnell, and Sebastian Walcher. On the quasi-steady-state approximation in an open Michaelis–Menten reaction mechanism. *AIMS Math*, 6:6781—-6814, 2021. ISSN 2473-6988.

[46] Justin Eilertsen, Malgorzata Anna Tyczynska, and Santiago Schnell. Hunting ε: The origin and validity of quasi-steady-state reductions in enzyme kinetics. *SIAM Journal on Applied Dynamical Systems*, 20(4):2450–2481, 2021.

[47] Justin Eilertsen, Santiago Schnell, and Sebastian Walcher. On the anti-quasi-steady-state conditions of enzyme kinetics. *Math. Biosci.*, 350:Paper No. 108870, 2022. ISSN 0025-5564,1879-3134.

[48] Justin Eilertsen, Kashvi Srivastava, and Santiago Schnell. Stochastic enzyme kinetics and the quasi-steady-state reductions: application of the slow scale linear noise approximation à la Fenichel. *J. Math. Biol.*, 85(1):Paper No. 3, 27, 2022. ISSN 0303-6812,1432-1416.

[49] Justin Eilertsen, Santiago Schnell, and Sebastian Walcher. Natural parameter conditions for singular perturbations of chemical and biochemical reaction networks. *Bull. Math. Biol.*, 85(6):Paper No. 48, 75, 2023. ISSN 0092-8240,1522-9602.

[50] Justin Eilertsen, Santiago Schnell, and Sebastian Walcher. The unreasonable effectiveness of the total quasi-steady state approximation, and its limitations. *J. Theor. Biol.*, 583:Paper No. 111770, 2024. ISSN 0022-5193,1095-8541.

[51] Justin Eilertsen, Santiago Schnell, and Sebastian Walcher. Rigorous estimates for the quasi-steady state approximation of the Michaelis-Menten reaction mechanism at low enzyme concentrations. *Nonlinear Anal. Real World Appl.*, 78:Paper No. 104088, 27, 2024. ISSN 1468-1218,1878-5719.

[52] Péter Érdi and János Tóth. *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*. Manchester University Press, 1989.

[53] I Feinberg, JD March, G Fein, TC Floyd, JM Walker, and L Price. Period and amplitude analysis of 0.5–3c/sec activity in NREM sleep of young adults. *Electroencephalography and Clinical Neurophysiology*, 44(2):202–213, 1978.

[54] Neil Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.*, 21:193–226, 1971/72.

[55] Neil Fenichel. Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations*, 31(1):53–98, 1979.

[56] Neil Fenichel and JK Moser. Persistence and smoothness of invariant manifolds for flows. *Indiana University Mathematics Journal*, 21(3):193–226, 1971.

[57] AJ Finnimore, M Roebuck, D Sajkov, and RD McEvoy. The effects of the GABA agonist, baclofen, on sleep and breathing. *European Respiratory Journal*, 8(2):230–234, 1995.

[58] Marcos G Frank and H Craig Heller. The Function(s) of Sleep. *Sleep-Wake Neurobiology and Pharmacology*, pages 3–34, 2019.

[59] Paul Franken, Irene Tobler, and Alexander A Borbély. Sleep homeostasis in the rat: simulation of the time course of EEG slow-wave activity. *Neuroscience Letters*, 130(2): 141–144, 1991.

[60] Simon J Fraser. The steady state and equilibrium approximations: A geometrical picture. *The Journal of Chemical Physics*, 88(8):4732–4738, 1988.

[61] Matilde Esther Fuentes, Ramón Varón, Manuela García-Moreno, and Edelmira Valero. Kinetics of intra-and intermolecular zymogen activation with formation of an enzyme–zymogen complex. *The FEBS Journal*, 272(1):85–96, 2005.

[62] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.

[63] Alexandra Goeke and Sebastian Walcher. A constructive approach to quasi-steady state reductions. *Journal of Mathematical Chemistry*, 52:2596–2626, 2014.

[64] Alexandra Goeke, Christian Schilli, Sebastian Walcher, and Eva Zerz. Computing quasi-steady state reductions. *Journal of Mathematical Chemistry*, 50:1495–1513, 2012.

[65] Alexandra Goeke, Sebastian Walcher, and Eva Zerz. Determining "small parameters" for quasi-steady state. *Journal of Differential Equations*, 259(3):1149–1180, 2015.

[66] Alexandra Goeke, Sebastian Walcher, and Eva Zerz. Classical quasi-steady state reduction—a mathematical characterization. *Physica D: Nonlinear Phenomena*, 345: 11–26, 2017.

[67] Martin Golubitsky and D Schaeffer. A theory for imperfect bifurcation via singularity theory. *Communications on Pure and Applied Mathematics*, 32(1):21–98, 1979.

[68] Pedro J Gonçalves, Jan-Matthis Lueckmann, Michael Deistler, Marcel Nonnenmacher, Kaan Öcal, Giacomo Bassetto, Chaitanya Chintaluri, William F Podlaski, Sara A Haddad, Tim P Vogels, et al. Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife*, 9:e56261, 2020.

[69] Johan Grasman. *Asymptotic methods for relaxation oscillations and applications*, volume 63 of *Applied Mathematical Sciences*. Springer New York, NY, 2012.

[70] Arthur Gretton, Karsten M Borgwardt, Malte J Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *The Journal of Machine Learning Research*, 13(1):723–773, 2012.

[71] F. G. Heineken, H. M. Tsuchiya, and R. Aris. On the mathematical status of the pseudo-steady hypothesis of biochemical kinetics. *Math. Biosci.*, 1:95–113, 1967.

[72] Geertje Hek. Geometric singular perturbation theory in biological practice. *J. Math. Biol.*, 60(3):347–386, 2010.

[73] Victor Henri. *Lois générales de l'action des diastases*. Librairie Scientifique A. Hermann, 1903.

[74] Narmada Herath and Domitilla Del Vecchio. Reduced linear noise approximation for biochemical reaction networks with time-scale separation: The stochastic tQSSA$^+$. *J. Chem. Phys.*, 148(9):094108, 2018.

[75] Heisuke Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero: II. *Annals of Mathematics*, 79(2):205–326, 1964.

[76] Heisuke Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero: I. *Annals of Mathematics*, 79(1):109–203, 1964. ISSN 0003486X, 19398980. URL http://www.jstor.org/stable/1970486.

[77] Geoffrey A Holdgate, Thomas D Meek, and Rachel L Grimley. Mechanistic enzymology in drug discovery: A fresh perspective. *Nature Reviews Drug Discovery*, 17(2):115–132, 2018.

[78] Nicholas HG Holford and Lewis B Sheiner. Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models. *Clinical Pharmacokinetics*, 6(6):429–453, 1981.

[79] Frank Charles Hoppensteadt. Singular perturbations on the infinite interval. *Transactions of the American Mathematical Society*, 123(2):521–535, 1966.

[80] John H. Hubbard and Beverly H. West. *Differential equations: A dynamical systems approach*, volume 5 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1995. ISBN 0-387-97286-2. Ordinary differential equations, Corrected reprint of the 1991 edition.

[81] Reto Huber, Tom Deboer, and Irene Tobler. Effects of sleep deprivation on sleep and sleep EEG in three mouse strains: empirical data and simulations. *Brain Research*, 857(1-2):8–19, 2000.

[82] Felix Hummel, Samuel Jelbart, and Christian Kuehn. Geometric blow-up of a dynamic Turing instability in the Swift-Hohenberg equation. *Journal of Differential Equations*, 427:219–309, 2025.

[83] Mark D Humphries, Jose Angel Obeso, and Jakob Kisbye Dreyer. Insights into Parkinson's disease from computational models of the basal ganglia. *Journal of Neurology, Neurosurgery & Psychiatry*, 89(11):1181–1188, 2018.

[84] John K Hunter. An introduction to real analysis. *Department of Mathematics, University of California at Davis*, 2014. URL https://www.math.ucdavis.edu/~hunter/intro_analysis_pdf/intro_analysis.pdf.

[85] Chandra L Jackson, Susan Redline, and Karen M Emmons. Sleep as a potential fundamental contributor to disparities in cardiovascular health. *Annual Review of Public Health*, 36(1):417–440, 2015.

[86] Hildeberto Jardón-Kojakhmetov and Christian Kuehn. A survey on the blow-up method for fast-slow systems. *arXiv preprint arXiv:1901.01402*, 2019.

[87] S Jelbart and C Kuehn. A formal geometric blow-up method for pattern forming systems. *Topics in Multiple Time Scale Dynamics, M. Engel, H. Jardón-Kojakhmetov, and C. Soresina, eds*, 806:49–86, 2024.

[88] Barbara E Jones. Arousal systems. *Frontiers in Bioscience*, 8(5):438–51, 2003.

[89] Christopher K. R. T. Jones. *Geometric singular perturbation theory*, pages 44–118. Springer Berlin Heidelberg, 1995. ISBN 978-3-540-49415-7. doi: 10.1007/BFb0095239. URL https://doi.org/10.1007/BFb0095239.

[90] Ilenna Simone Jones and Konrad Paul Kording. Efficient optimization of ODE neuron models using gradient descent. *arXiv preprint arXiv:2407.04025*, 2024.

[91] N.G. Van Kampen. Chapter X. The expansion of the Master Equation. In *Stochastic Processes in Physics and Chemistry (3rd Edition)*, North-Holland Personal Library, pages 244–272. Elsevier, Amsterdam, 2007.

[92] Xingye Kan, Chang Hyeong Lee, and Hans G. Othmer. A multi-time-scale analysis of chemical reaction networks: II. Stochastic systems. *J. Math. Biol.*, 73(5):1081–1129, 2016.

[93] Hye-Won Kang and Thomas G. Kurtz. Separation of time-scales and model reduction for stochastic reaction networks. *Ann. Appl. Probab.*, 23(2):529–583, 2013.

[94] Hye-Won Kang, Wasiur R. KhudaBukhsh, Heinz Koeppl, and Grzegorz Rempala. Quasi-steady-state approximations derived from the stochastic model of enzyme kinetics. *Bull. Math. Biol.*, 81:1303–1336, 2019.

[95] Hans G Kaper and Tasso J Kaper. Asymptotic analysis of two reduction methods for systems of chemical reactions. *Physica D*, 165(1-2):66–93, 2002.

[96] Robert E Kass, Valérie Ventura, and Emery N Brown. Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology*, 94(1):8–25, 2005.

[97] G. S. Katzenberger. Solutions of a stochastic differential equation forced onto a manifold by a large drift. *Ann. Probab.*, 19:1587–1628, 1991.

[98] Jae Kyoung Kim and John J. Tyson. Misuse of the Michaelis—Menten rate law for protein interaction networks and its remedy. *PLoS Comp. Biol.*, 16(10):1–21, 10 2020.

[99] Jae Kyoung Kim, Krešimir Josić, and Matthew R. Bennett. The validity of quasi-steady-state approximations in discrete stochastic simulations. *Biophys. J.*, 107(3):783 – 793, 2014.

[100] Jae Kyoung Kim, Krešimir Josić, and Matthew R Bennett. The relationship between stochastic and deterministic quasi-steady state approximations. *BMC Systems Biology*, 9(1):1–13, 2015.

[101] Jae Kyoung Kim, Grzegorz A. Rempala, and Hye-Won Kang. Reduction for Stochastic Biochemical Reaction Networks with Multiscale Conservations. *Multiscale Modeling & Simulation*, 15(4):1376–1403, 2017.

[102] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[103] Martin Krupa and Peter Szmolyan. Extending slow manifolds near transcritical and pitchfork singularities. *Nonlinearity*, 14(6):1473, 2001.

[104] Martin Krupa and Peter Szmolyan. Extending geometric singular perturbation theory to nonhyperbolic points—fold and canard points in two dimensions. *SIAM Journal on Mathematical Analysis*, 33(2):286–314, 2001.

[105] Martin Krupa and Peter Szmolyan. Relaxation oscillation and canard explosion. *Journal of Differential Equations*, 174(2):312–368, 2001.

[106] Christian Kuehn. *Multiple Time Scale Dynamics*, volume 191 of *Applied Mathematical Sciences*. Springer, 2015. ISBN 978-3-319-34417-1.

[107] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[108] Ajit Kumar and Krešimir Josić. Reduced models of networks of coupled enzymatic reactions. *Journal of Theoretical Biology*, 278(1):87–106, 2011.

[109] Christian Lax and Sebastian Walcher. Singular perturbations and scaling. *Discrete Contin. Dyn. Syst. Ser. B*, 25(1):1–29, 2020. ISSN 1531-3492,1553-524X. doi: 10.3934/dcdsb.2019170. URL https://doi.org/10.3934/dcdsb.2019170.

[110] Shi-Bin Li, Valentina Martinez Damonte, Chong Chen, Gordon X Wang, Justus M Kebschull, Hiroshi Yamaguchi, Wen-Jie Bian, Carolin Purmann, Reenal Pattni, Alexander Eckehart Urban, et al. Hyperexcitable arousal circuits drive sleep instability during aging. *Science*, 375(6583):eabh3021, 2022.

[111] Tong Li and Ming Yuan. On the optimality of Gaussian kernel based nonparametric tests against smooth alternatives. *Journal of Machine Learning Research*, 25(334): 1–62, 2024.

[112] Pierre-Hervé Luppi, Christelle Peyron, and Patrice Fort. Not a single but multiple populations of GABAergic neurons control sleep. *Sleep Medicine Reviews*, 32:85–94, 2017.

[113] Shev MacNamara, Alberto M. Bersani, Kevin Burrage, and Roger B. Sidje. Stochastic chemical kinetics and the total quasi-steady-state assumption: Application to the stochastic simulation algorithm and chemical master equation. *J. Chem. Phys.*, 129 (9):095105, 2008.

[114] Thomas Maiwald and Jens Timmer. Dynamical modeling and multi-experiment fitting with Potterswheel. *Bioinformatics*, 24(18):2037–2043, 2008.

[115] MC Manjabacas, E Valero, M García-Moreno, C Garrido, and R Varón. Kinetics of an autocatalytic zymogen reaction in the presence of an inhibitor coupled to a monitoring reaction. *Bulletin of Mathematical Biology*, 58:19–41, 1996.

[116] Henry B Mann and Donald R Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, pages 50–60, 1947.

[117] Ethan A. Mastny, Eric L. Haseltine, and James B. Rawlings. Two classes of quasi-steady-state model reductions for stochastic kinetics. *J. Chem. Phys.*, 127(9):094106, 2007.

[118] Leonor Michaelis, Maud L Menten, et al. Die kinetik der invertinwirkung. *Biochem. z*, 49(333-369):352, 1913.

[119] Alexandra Milik and Peter Szmolyan. Multiple time scales and canards in a chemical oscillator. In Christopher K. R. T. Jones and Alexander I. Khibnik, editors, *Multiple-Time-Scale Dynamical Systems*, pages 117–140, New York, NY, 2001. Springer New York. ISBN 978-1-4613-0117-2.

[120] E Mishchenko. *Differential equations with small parameters and relaxation oscillations*, volume 13 of *Mathematical Concepts and Methods in Science and Engineering*. Springer New York, NY, 2013.

[121] Ralph E Mistlberger. Circadian regulation of sleep in mammals: role of the suprachiasmatic nucleus. *Brain Research Reviews*, 49(3):429–454, 2005.

[122] Robert Y Moore. Suprachiasmatic nucleus in sleep–wake regulation. *Sleep Medicine*, 8:27–33, 2007.

[123] S Morairty, D Rainnie, R McCarley, and R Greene. Disinhibition of ventrolateral preoptic area sleep-active neurons by adenosine: A new mechanism for sleep promotion. *Neuroscience*, 123(2):451–457, 2004.

[124] An Hoang Nguyen and Simon J Fraser. Geometrical picture of reaction in enzyme kinetics. *J. Chem. Phys.*, 91:186–193, 1989.

[125] Lena Noethen and Sebastian Walcher. Quasi-steady state in the Michaelis–Menten system. *Nonlinear Anal. Real World Appl.*, 8(5):1512–1535, 2007.

[126] Lena Noethen and Sebastian Walcher. Quasi-steady state and nearly invariant sets. *SIAM Journal on Applied Mathematics*, 70(4):1341–1363, 2009.

[127] Lena Noethen and Sebastian Walcher. Tikhonov's theorem and quasi-steady state. *Discrete Contin. Dyn. Syst. Ser. B*, 16:945–961, 2011.

[128] Umberto Olcese, Steve K Esser, and Giulio Tononi. Sleep and synaptic renormalization: a computational study. *Journal of Neurophysiology*, 104(6):3476–3493, 2010.

[129] Chetan D. Pahlajani, Paul J. Atzberger, and Mustafa Khammash. Stochastic reduction method for biological chemical kinetics using time-scale separation. *J. Theo. Biol.*, 272 (1):96–112, 2011.

[130] Bernhard O. Palsson and Edwin N. Lightfoot. Mathematical modelling of dynamics and control in metabolic networks. I. On Michaelis-Menten kinetics. *J. Theor. Biol.*, 111:273–302, 1984.

[131] Sung-Ho Park, Justin Baik, Jiso Hong, Hanna Antila, Benjamin Kurland, Shinjae Chung, and Franz Weber. A probabilistic model for the ultradian timing of REM sleep in mice. *PLoS Computational Biology*, 17(8):e1009316, 2021.

[132] Todd L Parsons and Tim Rogers. Dimension reduction for stochastic dynamical systems forced onto a manifold by large drift: a constructive approach with examples from theoretical biology. *J. Phys. A Math. Theor.*, 50(41):415601, 2017.

[133] Morten Gram Pedersen, Alberto Maria Bersani, and Enrico Bersani. Quasi steady-state approximations in complex intracellular signal transduction networks–a word of caution. *Journal of Mathematical Chemistry*, 43(4):1318–1344, 2008.

[134] Morten Gram Pedersena, Alberto M Bersanib, and Enrico Bersanic. The total quasi-steady-state approximation for fully competitive enzyme reactions. *Bulletin of Mathematical Biology*, 69(1):433–457, 2007.

[135] Adrien Peyrache, Nima Dehghani, Emad N Eskandar, Joseph R Madsen, William S Anderson, Jacob A Donoghue, Leigh R Hochberg, Eric Halgren, Sydney S Cash, and Alain Destexhe. Spatiotemporal dynamics of neocortical excitation and inhibition

during human sleep. *Proceedings of the National Academy of Sciences*, 109(5):1731–1736, 2012.

[136] AJK Phillips, PA Robinson, and EB Klerman. Arousal state feedback as a potential physiological generator of the ultradian REM/NREM sleep cycle. *Journal of Theoretical Biology*, 319:75–87, 2013.

[137] Andrew JK Phillips, Charles A Czeisler, and Elizabeth B Klerman. Revisiting spontaneous internal desynchrony using a quantitative model of sleep physiology. *Journal of Biological Rhythms*, 26(5):441–453, 2011.

[138] Andrew JK Phillips, Ben D Fulcher, Peter Alexander Robinson, and Elizabeth B Klerman. Mammalian rest/activity patterns explained by physiologically based modeling. *PLOS Computational Biology*, 9(9):e1003213, 2013.

[139] Thomas D Pollard. A guide to simple and informative binding assays. *Molecular Biology of the Cell*, 21(23):4061–4067, 2010.

[140] Tarja Porkka-Heiskanen, Robert E Strecker, Mahesh Thakkar, Alvhild A Bjørkum, Robert W Greene, and Robert W McCarley. Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness. *Science*, 276(5316):1265–1268, 1997.

[141] Tarja Porkka-Heiskanen, RE Strecker, and RW McCarley. Brain site-specificity of extracellular adenosine concentration changes during sleep deprivation and spontaneous sleep: an in vivo microdialysis study. *Neuroscience*, 99(3):507–517, 2000.

[142] Ronald B Postuma, Alex Iranzo, Michele Hu, Birgit Högl, Bradley F Boeve, Raffaele Manni, Wolfgang H Oertel, Isabelle Arnulf, Luigi Ferini-Strambi, Monica Puligheddu, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain*, 142(3):744–759, 2019.

[143] Aaditya Ramdas, Nicolás García Trillos, and Marco Cuturi. On Wasserstein two-sample testing and related families of nonparametric tests. *Entropy*, 19(2):47, 2017.

[144] Jim O Ramsay, Giles Hooker, David Campbell, and Jiguo Cao. Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(5):741–796, 2007.

[145] JG Reich and EE Sel'kov. Mathematical-analysis of metabolis networks. *FEBS Lett.*, 43:S119–S127, 1974.

[146] Michael J Rempe, Janet Best, and David Terman. A mathematical model of the sleep/wake cycle. *Journal of Mathematical Biology*, 60(5):615–644, 2010.

[147] Michael J Rempe, Janne Grønli, Torhild Thue Pedersen, Jelena Mrdalj, Andrea Marti, Peter Meerlo, and Jonathan P Wisor. Mathematical modeling of sleep state dynamics in a rodent model of shift work. *Neurobiology of Sleep and Circadian Rhythms*, 5:37–51, 2018.

[148] Yuri Robbers, Eva AS Koster, Doortje I Krijbolder, Amanda Ruijs, Sander Van Berloo, and Johanna H Meijer. Temporal behaviour profiles of Mus musculus in nature are affected by population activity. *Physiology & Behavior*, 139:351–360, 2015.

[149] Johann Rudi, Julie Bessac, and Amanda Lenzi. Parameter estimation with dense and convolutional neural networks applied to the Fitzhugh–Nagumo ODE. In *Mathematical and Scientific Machine Learning*, pages 781–808. PMLR, 2022.

[150] Thomas Rusterholz, Roland Dürr, and Peter Achermann. Inter-individual differences in the dynamics of sleep homeostasis. *Sleep*, 33(4):491–498, 2010.

[151] Soheil Saghafi, Timothy Rumbell, Viatcheslav Gurev, James Kozloski, Francesco Tamagnini, Kyle CA Wedgwood, and Casey O Diekman. Inferring parameters of pyramidal neuron excitability in mouse models of Alzheimer's disease using biophysical modeling and deep learning. *Bulletin of Mathematical Biology*, 86(5):46, 2024.

[152] K.R. Sanft, D. T. Gillespie, and L. R. Petzold. The legitimacy of the stochastic Michaelis–Menten approximation. *IET Syst. Biol.*, 5:58–69, 2011.

[153] Clifford B Saper and Patrick M Fuller. Wake–sleep circuitry: an overview. *Current Opinion in Neurobiology*, 44:186–192, 2017.

[154] Clifford B Saper, Thomas C Chou, and Thomas E Scammell. The sleep switch: hypothalamic control of sleep and wakefulness. *Trends in Neurosciences*, 24(12):726–731, 2001.

[155] Clifford B Saper, Thomas E Scammell, and Jun Lu. Hypothalamic regulation of sleep and circadian rhythms. *Nature*, 437(7063):1257–1263, 2005.

[156] Nicolae Sapoval, Amirali Aghazadeh, Michael G Nute, Dinler A Antunes, Advait Balaji, Richard Baraniuk, C_J Barberan, Ruth Dannenfelser, Chen Dun, Mohammadamin Edrisi, et al. Current progress and open challenges for applying deep learning across the biosciences. *Nature Communications*, 13(1):1728, 2022.

[157] Thomas E Scammell, Elda Arrigoni, and Jonathan O Lipton. Neural circuitry of wakefulness and sleep. *Neuron*, 93(4):747–765, 2017.

[158] S. Schnell and P. K. Maini. A century of enzyme kinetics. Reliability of the K_M and $v_{\max}$ estimates. *Comments Theor. Biol.*, 8:169–187, 2003.

[159] S Schnell and C Mendoza. Closed form solution for time-dependent enzyme kinetics. *Journal of Theoretical Biology*, 187(2):207–212, 1997.

[160] Santiago Schnell. Validity of the Michaelis–Menten equation – steady-state or reactant stationary assumption: that is the question. *FEBS J.*, 281(2):464–472, 2014.

[161] Santiago Schnell and Philip K. Maini. Enzyme kinetics at high enzyme concentration. *Bull. Math. Biol.*, 62:483–499, 2000.

[162] Santiago Schnell and Philip K Maini. Enzyme kinetics at high enzyme concentration. *Bulletin of Mathematical Biology*, 62:483–499, 2000.

[163] L. A. Segel and M. Slemrod. The quasi-steady-state assumption: A case study in perturbation. *SIAM Review*, 31:446–477, 1989.

[164] Lee A. Segel. On the validity of the steady state assumption of enzyme kinetics. *Bull. Math. Biol.*, 50:579–593, 1988.

[165] Jerome M Siegel. Clues to the functions of mammalian sleep. *Nature*, 437(7063): 1264–1271, 2005.

[166] Anne C Skeldon, Gianne Derks, and Derk-Jan Dijk. Modelling changes in sleep timing and duration across the lifespan: Changes in circadian rhythmicity or sleep homeostasis? *Sleep Medicine Reviews*, 28:96–107, 2016.

[167] Kashvi Srivastava, Justin Eilertsen, Victoria Booth, and Santiago Schnell. Accuracy versus Predominance: Reassessing the validity of the quasi-steady-state approximation. *Bulletin of Mathematical Biology*, 87(6):73, 2025.

[168] Wylie Stroberg and Santiago Schnell. On the estimation errors of k_m and v from time-course experiments using the Michaelis–Menten equation. *Biophysical Chemistry*, 219: 17–27, 2016.

[169] Peter Szmolyan and Martin Wechselberger. Canards in R3. *Journal of Differential Equations*, 177(2):419–453, 2001.

[170] Philipp Thomas, Arthur V. Straube, and Ramon Grima. Communication: Limitations of the stochastic quasi-steady-state approximation in open biochemical reaction networks. *J. Chem. Phys.*, 135(18):181103, 2011.

[171] Philipp Thomas, Ramon Grima, and Arthur V. Straube. Rigorous elimination of fast stochastic variables from the linear noise approximation using projection operators. *Phys. Rev. E*, 86:041110, 2012.

[172] Philipp Thomas, Arthur V. Straube, and Ramon Grima. The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. *BMC Sys. Biol.*, 6(1):39, 2012.

[173] A.N. Tikhonov. Systems of differential equations containing small parameters in their derivatives. *Mat. Sb. (N.S.)*, 31:575–586, 1952.

[174] Andrei Nikolaevich Tikhonov. Systems of differential equations containing small parameters in the derivatives. *Matematicheskii sbornik*, 73(3):575–586, 1952.

[175] Nicholas Tolley, Pedro LC Rodrigues, Alexandre Gramfort, and Stephanie R Jones. Methods and considerations for estimating parameters in biophysically detailed neural models with simulation based inference. *PLOS Computational Biology*, 20(2):e1011108, 2024.

[176] Lorenz Trachsel, Irene Tobler, Peter Achermann, and Alexander A Borbély. Sleep continuity and the REM-nonREM cycle in the rat under baseline conditions and after sleep deprivation. *Physiology & Behavior*, 49(3):575–580, 1991.

[177] Abraham R Tzafriri and Elazer R Edelman. The total quasi-steady-state approximation is valid for reversible enzyme kinetics. *Journal of Theoretical Biology*, 226(3): 303–313, 2004.

[178] Gloria I Valderrama-Bahamóndez and Holger Fröhlich. MCMC techniques for parameter estimation of ODE based models in systems biology. *Frontiers in Applied Mathematics and Statistics*, 5:55, 2019.

[179] Eve Van Cauter, Laurence Plat, Martin B Scharf, Rachel Leproult, Sonya Cespedes, Mireille L'Hermite-Baleriaux, Georges Copinschi, et al. Simultaneous stimulation of slow-wave sleep and growth hormone secretion by gamma-hydroxybutyrate in normal young men. *The Journal of Clinical Investigation*, 100(3):745–753, 1997.

[180] Ferdinand Verhulst. *Methods and applications of singular perturbations: boundary layers and multiple timescale dynamics*, volume 50. Springer Science & Business Media, 2006.

[181] Julie Vienne, Bernhard Bettler, Paul Franken, and Mehdi Tafti. Differential effects of GABAB receptor subtypes, γ-hydroxybutyric acid, and baclofen on EEG activity and sleep regulation. *Journal of Neuroscience*, 30(42):14194–14204, 2010.

[182] VV Vyazovskiy, P Achermann, and I Tobler. Sleep homeostasis in the rat in the light and dark period. *Brain research bulletin*, 74(1-3):37–44, 2007.

[183] Franz Weber and Yang Dan. Circuit-based interrogation of sleep control. *Nature*, 538 (7623):51–59, 2016.

[184] Martin Wechselberger. *Geometric Singular Perturbation Theory Beyond the Standard Forms*. Number 6 in Frontiers in Applied Dynamical Systems: Tutorials and Reviews. Springer, 2020.

[185] Fabrice Wendling. Neurocomputational models in the study of epileptic phenomena. *Journal of Clinical Neurophysiology*, 22(5):285–287, 2005.

[186] ET Whittaker and GN Watson. *A Course of Modern Analysis, (3rd Edition)*. Cambridge Mathematical Library. Cambridge University Press, 1920.

[187] Jia-Wei Wu, Ying Wu, and Zhi-Xin Wang. Kinetic analysis of a simplified scheme of autocatalytic zymogen activation. *European Journal of Biochemistry*, 268(6):1547–1553, 2001.

[188] Choujun Zhan and Lam F Yeung. Parameter estimation in systems biology models using spline approximation. *BMC Systems Biology*, 5:1–12, 2011.

[189] Ji Zhao and Deyu Meng. FastMMD: Ensemble of circular discrepancy for efficient two-sample test. *Neural Computation*, 27(6):1345–1372, 2015.